ESTABLISHMENT, CHARACTERISATION AND MiRNA TRANSCRIPTOME PROFILING OF SPHEROID-ENRICHED CELLS WITH BREAST CANCER STEM CELL PROPERTIES

By

BOO LILY

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

Faculty of Medicine and Health Sciences,

Universiti Tunku Abdul Rahman,

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy of Medical Sciences

June 2020

ABSTRACT

ESTABLISHMENT, CHARACTERISATION AND MiRNA TRANSCRIPTOME PROFILING OF SPHEROID-ENRICHED CELLS WITH BREAST CANCER STEM CELL PROPERTIES

Boo Lily

Cancer stem cells (CSCs) are self-renewing cancer cells and are thought to be a source of tumour recurrence. The CSCs population could be enriched in serum-free culture condition as this environment favoured their expansion while the rest of non-CSCs population undergo anoikis. MCF-7 cells, which being the luminal type are non-metastatic, and MDA-MD-231 cells, which are negative for the three breast receptors and regarded as highly aggressive were used in this study. MicroRNAs (miRNAs) regulate both normal stem cells and CSCs, and deregulation of miRNAs has an important role of tumorigenesis. Although there are already some miRNAs that have been reported in breast cancer, precise information on miRNAs involved in breast CSCs is still lacking. Therefore, we sought to evaluate the phenotypic characteristics of the spheroid-enriched cells for their CSCs properties and also to determine the miRNA expression profile. We confirmed the enrichment of the spheroid-enriched cancer stem cells-like from human breast cancer cell lines, MCF-7 and MDA-MB-231 by evaluating the characteristics of the *in vitro* spheroid-enriched cells. After obtaining the spheroid-enriched cells, miRNA next generation sequencing and real-time polymerase chain reaction were performed to evaluate their miRNA profile. Our results showed that the spheroid cells

derived from both breast cancer cell lines were enriched with CSCs properties namely self-renewability, expression of stem cells-related markers, and enhancement of drug resistance. Using a 2-fold expression as the cut-off point, 25 up-regulated and 97 down-regulated differentially expressed miRNAs were identified in MCF-7 spheroid cells compared to their parental cells. On the other hand, 30 up-regulated and 36 down-regulated differentially expressed miRNAs were found in MDA-MB-231 spheroid in relative to the parental cells. The targeted genes from the uniquely deregulated miRNAs found in MCF-7 spheroid cells were enriched in pathways involved in induction of epithelial-mesenchymal transition (EMT) responsible for enhancing invasion and migration of the cells. On the other hand, the targeted genes from the uniquely deregulated miRNAs found in MDA-MB-231 spheroid cells were enriched in pathways involved in maintaining EMT associated with the increased chemoresistance. A total of 20 miRNAs including miR-15b, miR-34a, miR-148a, miR-196b, and miR-628 were found to be commonly deregulated between these two breast cancer spheroid-enriched CSCs cell types, which highlights the involvement of these miRNAs in maintaining the CSCs features. The enriched genes were involved in core pathways found in stem cells primarily on focal adhesion, MAPK, Wnt, Hedhehog, mTOR and VEGF. The levels of the selected miRNAs measured by real-time PCR and NGS showed similar trend, indicating the reliability of the sequencing data and supporting the interpretation of the expression profiles and pathways information based on the miRNAs expression results in this study. We have demonstrated that the spheroid culturing method can be used to enrich for CSCs-like subpopulations in both breast cancer cell lines as shown in the

present study. Our data suggest that there were distinct miRNA expression profiles in spheroid relative to parental cells for both breast cancer cell lines. This reflects that the phenotypic behaviour and other distinctive features of spheroid-enriched CSCs in MCF-7 and MDA-MB-231 are regulated by miRNAs. Further studies are needed to validate whether the panel of these distinct miRNAs could be used as potential molecular targets for clinical applications of breast cancer stem cells.

ACKNOWLEDGEMENTS

First and foremost, I am grateful to the Lord for establishing me to complete this thesis. My special thanks first go to my supervisor, Professor Dr. Alan Ong Han Kiat, for his support, guidance, patience throughout my whole study. I would also like to express my gratitude to my co-supervisor, Emeritus Professor Dr. Cheong Soon Keng for his continuous guidance and support throughout my studies. Special thanks to my external co-supervisor, Associate Professor Dr.Yeap Swee Keong, for his invaluable suggestions to my projects during my graduate study. Without their guidance and persistent help, this thesis would not have been possible.

I would also like to thank Faculty of Medicine and Health Sciences, Universiti Tunku Abdul Rahman and Institute of Bioscience, Universiti Putra Malaysia for providing me with all the necessary facilities. This work was supported by the Ministry of Science, Technology and Innovation of Malaysia (MOSTI) for the E-Science Fund (02-02-11-SF0125), UM High Impact Research Grants (UM.C/625/1/HIR/MOHE/CHAN/01, A000001-50001, UM.C/625/1/HIR/MOHE/CHAN/14/1, H-50001-A000027) and UTAR Research Fund (UTARRF A6200/A23). Special thanks to MyBrain15 Scholarship, Ministry of Higher Learning Malaysia for the opportunity to finance my education. Receiving this scholarship help reduce my financial burden as I continue to pursue my tertiary education.

In addition, I would like to thank my team members for their discussions and technical efforts which have been very valuable on my research. They are Dr. Ho Wan Yong, Dr. Huynh Ky, Dr. Dilan Amila Satharasinghe, Ms. Nancy Liew and Dr. Tan Sheau Wei. Special mention to Ms. Norlaily Mohd. Ali for her encouragement and support. My dear friend, *Dora*, thank you for always be there and giving me the extra push I need when I was at the verge of giving up. In addition, I would also like to express my gratitude to all the lab members in UTAR FMHS Postgraduate Laboratories, including Dr. Choong Pei Feng (Erica), Ms. Teh Hui Xin, and Mr. Lim Sheng Jye. It was rewarding, enjoyable and fun to work along with you all.

Last but certainly not least, my deepest gratitude to my family for their unconditional love and constant encouragement in my pursuit of doctorate degree. My dear parents, Pa and Ma, for their selfless love and having faith in me to see me through one of the most difficult journey in my life. To my Ma especially, words could not describe how grateful I am to have you. I salute you for all the sacrifices that you made during this challenging period. I am also indebted to my two elder sisters Jannis and Chinie for their support and valuable prayers. And finally, to my husband, Woon Chuan for his continued understanding, unfailing love and support during my pursuit. And to my two little princesses, Emily and Beverly, thanks for being such good babies during this journey.

Again, thanks to everyone for believing in me for all these years, shower me with strength and patience so that I can complete what I started.

APPROVAL SHEET

This thesis entitled "**ESTABLISHMENT, CHARACTERISATION AND MiRNA TRANSCRIPTOME PROFILING OF SPHEROID-ENRICHED CELLS WITH BREAST CANCER STEM CELL PROPERTIES"** was prepared by BOO LILY and submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy of Medical Sciences at Universiti Tunku Abdul Rahman.

Approved by:

_____

(Prof. Dr. Alan Ong Han Kiat)

Date: …….....

Professor/Supervisor

Department of Pre-clinical Sciences

Faculty of Medicine and Health Sciences

Universiti Tunku Abdul Rahman

_____

(Emeritus Prof.Dr. Cheong Soon Keng)

Date: …….....

Senior Professor/Co-supervisor

Department of Medicine

Faculty of Medicine and Health Sciences

Universiti Tunku Abdul Rahman

FACULTY OF MEDICINE AND HEALTH SCIENCES

UNIVERSITI TUNKU ABDUL RAHMAN

Date: ____ _____

SUBMISSION OF THESIS

It is hereby certified that **Boo Lily** (ID No: **14UMD07972**) has completed this thesis entitled "**ESTABLISHMENT, CHARACTERISATION AND MiRNA TRANSCRIPTOME PROFILING OF SPHEROID-ENRICHED CELLS WITH BREAST CANCER STEM CELL PROPERTIES"** under the supervision of Prof. Dr. Alan Ong Han Kiat (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, Universiti Tunku Abdul Rahman.

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,

_____

(Boo Lily)

DECLARATION

I BOO LILY 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.

_____

BOO LILY

Date _____

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	2-dimensional
3D	3-dimensional
5-FU	5-fluorouracil
AB	Alamar blue
ABC	ATP binding cassette
ALDH	Aldehyde dehydrogenase enzyme
AML	Acute myeloid leukaemia
ATCC	American Type Culture Collection
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
Bcl-2	B-cell lymphoma 2
BP	Biological processes
BrCa	Breast cancer
BRCA1	Breast cancer type 1 susceptibility protein
BRCA2	Breast cancer type 2 susceptibility protein
CC	Cellular components
CD133	Cluster of differentiation 133
CD24	Cluster of differentiation 24
CD34	Cluster of differentiation 34
CD38	Cluster of differentiation 38
CD44	Cluster of differentiation 44
CD49f	Cluster of differentiation 49f
cDNA	Complementary deoxyribonucleic acid
CLC	Genomics Workbench
CO2	Carbon dioxide
CSCs	Cancer stem cells
Ct	threshold cycle
DAPI	4,6-diamidino-2-phenylindole
DAVID	Database for Annotation, Visualization and Integrated Discovery
DE	Differentially expressed
DEAB	Diethylaminobenzaldehyde

DGCR8	DiGeorge Syndrome Critical Region 8
DICER	Endoribonuclease
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphate
E2F1	E2F family of transcription factors
EGF	Epidermal growth factors
EGRF	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
ENPEP	Glutamyl aminopeptidase
EpCAM	Epithelial cell adhesion molecule
ER	Oestrogen receptors
FBS	Foetal bovine serum
FC	Fold change
FDR	False discovery rate
FGF	Fibroblast growth factors
FITC	Fluorescein isothiocyanate
FTP	File Transfer Protocol
FUT4	Fucosyltransferase 4
G0	G zero phase
G1	G one phase
G2	G two phase
GAS1	Growth arrest-specific 1
GEO	Gene Expression Omnibus
GO	Gene ontology
HCl	Hydrochloric acid
HER2	Human epidermal growth factor receptor 2
HIF	Hypoxia-inducible factor
HMGA2	High-mobility group AT-hook 2
IC_{50}	Inhibitory concentration of 50%
ID4	DNA-binding protein inhibitor 4
IGF	Insulin-like growth factors

IL-11	Interleukin 11
ITCH	Ubiquitin-conjugating enzyme.
JAK-STAT	Janus Kinase/Signal Transducer and Activator of Transcription
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KLF4	Kruppel-like factor 4
LAMC1	Laminin subunit gamma 1
LNA	Locked nucleic acid
LRP2	Lipoprotein-related protein 2
LSCs	Leukemic stem cells
MAPK	Mitogen-Activated Protein Kinase Kinase
MCF-7	Oestrogen-dependent human breast cancer cell line
MDA-MB-231	Oestrogen-independent human breast cancer cell line.
MF	Molecular functions
miRNA	microRNA
mRNA	Messenger RNA
mTOR	Mammalian Target of Rapamycin
MTT	4,5-dimethylthiazol-2-yl
NaOAc	Sodium acetate
NCBI	National Center for Biotechnology Information
NGS	Next generation sequencing
NOD/SCID	Nonobese diabetic/severe combined immunodeficiency
NPC	Nasopharyngeal carcinoma
Oct 4	Octamer-binding transcription factor 4
P10	Passage 10
P5	Passage 5
p53	Cellular tumour antigen p53
PAGE	Polyacrylamide Gel
PARP	Poly ADP ribose polymerase
PAZ	Piwi/Argonaute/Zwille
PBS	Phosphate buffered saline
PCA	Principle component analysis
PCR	Polymerase Chain Reaction

PDGFR	Platelet Derived Growth Factor Receptor
PDT	Population doubling time
PE	Phycoerythrin
PR	Progesterone receptors
Q30	Quality score of 30
QC	Quality control
qRT	Quantitative reverse transcription
RIN	RNA integrity number
RISC	RNA interference silencing complex
RNA	Ribonucleic acid
RPMI-1640	Roswell Park Memorial Institute (RPMI-1640) cell culture media
rRNA	Ribosomal RNA
SALL4	Sal-like protein 4
SD	Standard deviation
SEM	Scanning electron microscope
SETDB1	SET domain bifurcated histone lysine methyltransferase 1
SFE	Sphere-forming efficiency
SMAD	Mothers against decapentaplegic homolog 4
Sox 2	Sex-determining region Y (SRY)-related box 2
SP	Side population
STAT3	Signal Transducer and Activator of Transcription 3
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TEMED	Tetramethylethylenediamine
TGFß	Transforming growth factor beta
TNBC	Triple negative breast cancer
TP53	Tumour protein 53
tRNA	Transfer ribonucleic acid
TS	Tumour-suppressive
TWF1	Twinfilin, actin-binding protein, homolog 1
TWIST1	Twist-related protein 1
UV	Ultraviolet

VEGF	Vascular endothelial growth factor
Wnt	Wingless-related integration site
ZEB	Zinc Finger E-Box Binding Homeobox 1

CHAPTER 1

INTRODUCTION

Globally, there are more than 200 different forms of cancer, accounting for about 1 in 6 deaths in 2018 (Siegel, Miller et al., 2019). Despite generations of extensive research on cancer onset, spread, prevention and effective treatment, there are still many uncertainties on the complexity of the pathogenesis of cancers (Thomas, Kelly et al., 2019). It has become clear that the accumulation of genetic damage caused by mutations lead to uncontrolled cell division of cancer cells (Basu, 2018). These abnormal cells can grow into mass of cells which can invade other parts of the body in a process known as metastasis (Welch and Hurst, 2019). Therefore, the search for the underlying key molecular mechanisms that causes the transformation of normal cells into cancerous cells is a priority.

Recent worldwide reports stated that the most common cancers are lung (12.3% of cases), followed by breast (12.3% of cases) and prostate (7.5% of cases) (Ferlay, Colombet et al., 2018). Breast cancer is the most common form of cancer in women and it has been predicted that in the United States alone, 1 in 8 women will develop breast cancer in her lifetime (Bray, Ferlay et al., 2018). Breast cancer caused about 1 in 30 of all deaths in Malaysia in 2016, and 88% of the reported deaths were avoidable (Asnarulkhadi, Maryam et al., 2016). These excess number of deaths recorded could be attributed due to late presentation (57%) and poor access to effective treatment strategies (43%) (Lum, 2019). Mortality from breast cancer remains high despite improvements

in surgery and common chemotherapies (Murallitharan, M., 2018). Cancer recurrence and distant metastasis accounted for 49% of survival rates as patients failed to respond to these treatments (Abdullah NA, Wan Mahiyuddin WR et al., 2013).

Emergence evidences have shown that there exists a subpopulation of cells within the tumours that drive the tumour progression and recurrence (Geurts, Witteveen et al., 2017). This group of cells termed as cancer stem cells (CSCs) have stem-cell like features with self-renewing capabilities and possess enhanced chemoresistance ability (Capp, 2019). Current treatment approaches targeted on the bulk tumour could not possibly eradicate the CSCs, allowing CSCs to self-renew and repopulate the whole tumour (Pan, Ma et al., 2018). Therefore, a better understanding on the biology and molecular mechanisms that govern CSCs could provide cure for breast cancer (Yoshida and Saya, 2016).

To model breast CSCs research in the laboratories, breast cancer cell lines have been widely used as an *in vitro* model (Ledford, 2016). In this study, the two most widely used and published breast cancer cell lines, namely, MCF-7 and MDA-MB-231, representing the two major clinically different types of breast cancer, were used (Dai, Cheng et al., 2017). Isolation of breast CSCs populations from the tumours remains a challenge due to the low existence of these subpopulations of the total tumour cells, which was also reported in other types of tumours (Bao, Ahmad et al., 2013, Dobbin and Landen, 2013). Spheroid culturing technique based on the ability of the cells to survive in a

low adherent serum-free condition has been used to enrich CSCs subpopulation (Lee, Sung et al., 2018, Ward Rashidi, Mehta et al., 2019). Though these spheroid models have been widely used for drug discovery studies, the biological and molecular characteristics of their CSCs properties are still lacking (De Angelis ML, Bruselles A et al., 2018, Ho, Yeap et al., 2012, Li, Zhang et al., 2010).

A small short non-coding RNA, known as microRNA (miRNA) has been recently identified to play important roles in regulating gene expressions (Catalanotto, Cogoni et al., 2016, Hanna, Hossain et al., 2019). Deregulation of these miRNAs contributes to various human diseases including cancer and more recently as a common feature in CSCs (Garg, 2015, Giza, Vasilescu et al., 2014, Li and Kowdley, 2012). Although the research on miRNAs is rapidly developing, there are limited studies available regarding miRNA expression profiles of breast CSCs. Based on their significant roles, establishment of miRNAs profiles in particularly looking at breast CSCs has drawn attention from investigators (Vahidian, Mohammadi et al., 2018, Wang, Chen et al., 2017). Detection of specific miRNAs signatures together with their associated regulatory pathways and their functions in breast CSCs could provide potential miRNA-biomarkers for early diagnosis and treatment (Cha, Choi et al., 2017, Jiang, Ma et al., 2019).

The present study hypothesised that the anchorage-independent culture technique in the form of spheroid could enrich cells with CSCs properties from the two distinct breast cancer subtypes. We also hope that the second focus of

our study on the miRNA profiling on CSCs subtypes could complement our understanding on the molecular mechanism at the miRNA level to facilitate the development of more effective targeted therapies.

CHAPTER 2

LITERATURE REVIEW

2.1 Cancer

Tumour is a disease caused by an abnormal appearance of mass of tissues in a body. It is caused by the uncontrolled growth of cells which evade the body's surveillance mechanism and continue to proliferate without stopping (Blasco, 2005). Tumours can be classified into two types, i. benign and ii. malignant (Hutchinson, 2010). Cells in benign tumour have limited growth ability and do not spread to nearby tissue. Malignant tumours, on the contrary, are more dangerous, as the cells grow out of control and invade parts of the body. The invasion occurred when the cells dislodge and break off from the initial tumour, form a secondary tumour and metastasise (spread) into other organs (Akrap, Andersson et al., 2016). A broader way to classify cancer cell types rather than the particular cancers is according to the site of cancer origin. The different types of cancers are known as carcinoma (refers to cancer originated from the external lining of organs and tracts), sarcoma (cancer originated in connective tissues such as bones, tendons and muscle), myeloma (refers to cancer arising from plasma cell of bone marrow), leukaemia (refers to cancer originated from blood forming tissues), lymphoma (refers to cancer arising from networks of lymphatic system) and brain and spinal cord cancers (cancers that affect the central nervous system). Although there are different types of cancer depending on the tissue of origins, they all share some common characteristics.

It has been proposed there are six important hallmarks that define what a cancer is, that simplify our understanding on the complexities of this disease (Hanahan and Weinberg, 2000). The six hallmarks are continuous growth, insensitivity to growth elimination, triggering growth invasion and metastasis, activation unlimited replicative ability, sustaining angiogenesis, and escaping apoptosis. Recently, the same research group has added two new hallmarks into the core list which are the reprogramming of energy metabolism and insensitivity to immune mechanism and two more emerging traits: genomic instability and mutation and tumour-promoting inflammatory responses (Hanahan and Weinberg, 2011). Altogether, these eight hallmarks encompass on the multistep processes on how normal cells progress to become cancerous.

Genetics changes and lifestyle factors are the main causes of cancer (Loef and Walach, 2012, Vogelstein, Papadopoulos et al., 2013). Cells in the body grow and reproduce in a coordinated way under strict regulations by the nucleus. However, exposure to carcinogens (cancer-causing substances) or inherited genetic changes cause the changes in genes that regulate the normal functioning of the cells. The process of change is known as mutation (Weinhold, Jacobsen et al., 2014). It is now evident cells with mutations in three main classes of genes such tumour suppressor genes, DNA repair genes and proto-oncogenes are the main contributors for cells to divide uncontrollably (Bailey, Tokheim et al., 2018). Alterations of the TP53 gene (well known as ‘guardian of the genome’), a gene that has been found in a wide range of human cancers, encodes proteins to arrest cell division and eventually blocks the formation of tumours (Giacomelli, Yang et al., 2018). Numerous

tumour-based p53 therapies have been described in recent years, with the main strategy of restoring the mutated p53 genes with healthy p53 genes in anti-cancer treatment (Aisner, Sholl et al., 2018, Liu, Zhang et al., 2015). Still, the potential applications of these treatments remain a challenge as it involves the use of genetically engineered virus delivery mechanisms that are normally associated with harmful side effects. Interestingly, tumour cells have been found to have an overexpression of oncogenes which are responsible for continuous cell proliferation, cell growth and differentiation (Liu, Murphy et al., 2018, Volonte, Vyas et al., 2018, Zhang, Yang et al., 2018). Despite decades of research on cancer, the molecular mechanisms of cancers which provide important opportunities for targeted cancer therapies are still poorly understood.

Cancers remain a difficult disease to treat and it is now a major problem across the world. It is the second most common cause of death after heart disease in the United States (Siegel, Miller et al., 2018). According to the global estimates from Institute for Health Metrics and Evaluation (IHME)'s global burden of diseases analysis, the attributed deaths due to cancer is on the increase, though many cases go unreported. Tracheal, bronchus and lung cancers have the highest mortality claiming 1.9 million in 2017 followed by colon, liver and breast cancers, all claiming deaths between 600,000 and 900,000 annually. The detailed breakdown on cancer deaths by type of cancer in the world in 2017 is shown in Figure 2.1. In Malaysia, there are about 100,000 cases of cancer reported each year, accounted for 13.6% of all deaths (Azizah Ab M, Nor Saleha I.T et al., 2016). According to the National Cancer

Patient Registry report 2007–2011 (Azizah Ab M, Nor Saleha I.T et al., 2016), cases of cancer reported were higher in females (54.8%) compared to males (45.2%), with life time risk of getting cancer reported to be 1 in 9 for females and 1 in 10 in males. The most common cancers affecting males and females in Malaysia are shown in Figure 2.2. It is projected that the number of cancers would increase by 15% by the year 2020, mainly affecting the older Malaysian population.

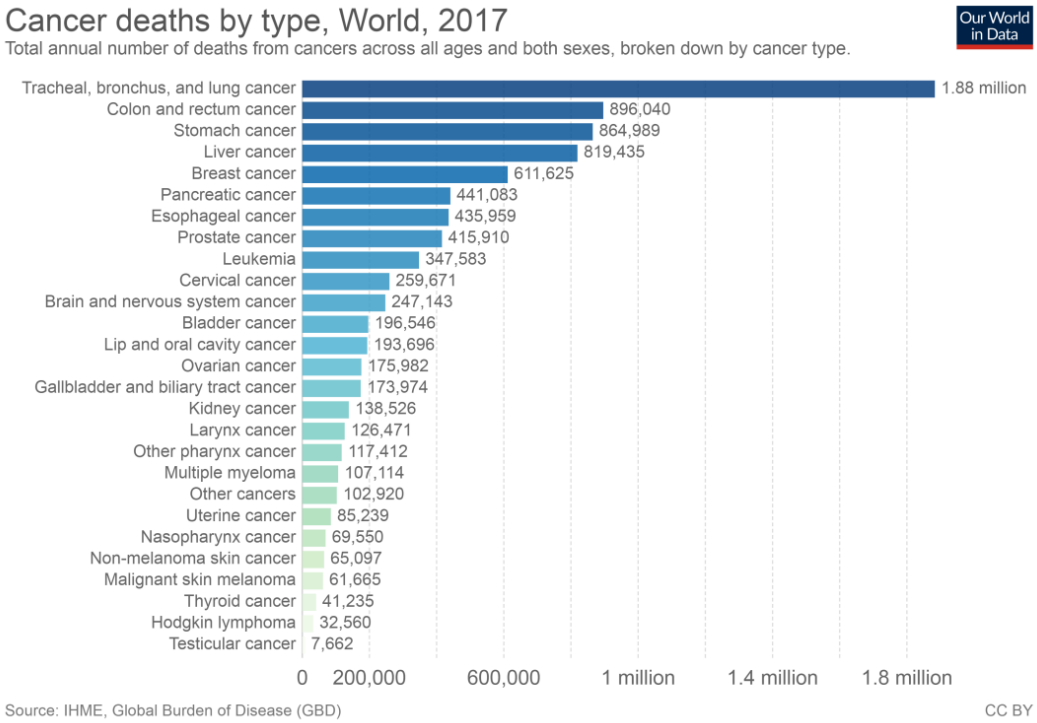


Figure 2.1. The bar charts show the detail breakdown on the worldwide pattern of cancer deaths. (Max and Hannah, 2018)

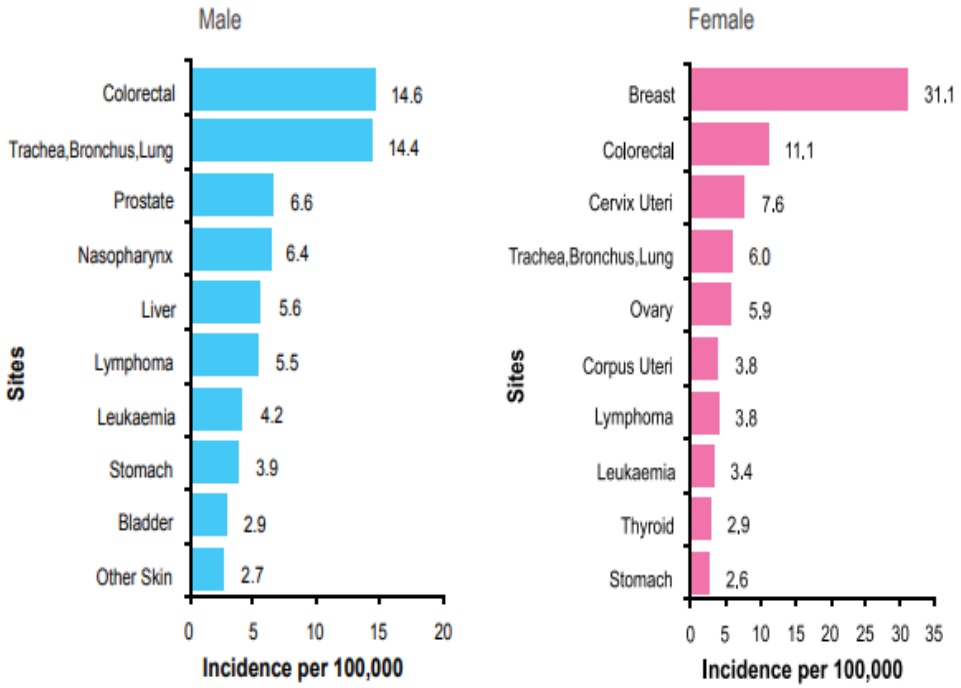


Figure 2.2. Age-standardised rate for ten common cancers by sex in Malaysia from 2007 to 2011 (Azizah Ab M, Nor Saleha I.T et al., 2016)

2.1.1 Global and Regional Burdens of Breast Cancer

Breast cancer is defined as the cancer which occurs in the tissues of the breast such as the lobes, lobules (milk glands), ducts, lymph nodes and also other parts of the breast (Feng, Spezia et al., 2018). It is commonly diagnosed in women worldwide with equal incidences rates in both developed and developing countries. In the United States alone, it is the second most common cause of death among women, in which 1 in 8 women would be diagnosed with breast cancer in her lifetime. Though it affects women generally, men somehow could develop breast cancer, but the occurrence is very rare. Across the globe, it is estimated that 1.6 million women are diagnosed for breast cancer yearly, accounting for 25% of all cancers in females (Torre, Islami et al., 2017).

In Malaysia, according to the National Cancer Registry, breast cancer incident rates displayed a slight variation with 1 in 20 women developing breast cancer in their lifetime, with different rates among the three main races. It was most common among the Chinese and Indian women (1 in 16) and 1 in 28 among the Malays population. Differences for this variation include demographic (geographic location, socio-economic), family history, reproductive factors (menarche age, first child, number of children), duration of breastfeeding, and lifestyle factors (diet and tobacco smoking) (Grosso, Bella et al., 2017).

2.1.2 Biology of Breast Cancer

Human breast tissues start to develop at the 6th week of life in the womb, leaving two breasts buds on the chest at birth. The breasts continue to develop further at puberty or pregnancy under the influence of female hormones specifically the oestrogen and progesterone (Javed and Lteif, 2013). The breast, known as the site of milk production, is surrounded by a mass of glands, connective, adipose, and fibrous tissues. The breast tissues are also made up of two types of networks, namely the blood system as well as the lymphatic system. The circulating blood in the breasts provides nutrient and oxygen to the breast cells and carries waste materials away from the breasts (Zucca-Matthes, Urban et al., 2016). Lymphatic systems on the other hand, drain excess fluid and provide immunity by killing foreign invasion and destroy cancerous cells. Breast development is largely under the influence of the oestrogens hormones, which peaks during puberty in females (Hilton, Clarke et al., 2018).

Oestrogens are steroid hormones released from the ovaries, responsible for the growth and development in female mammary gland and reproductive system (Wilkinson and Hardman, 2017). In addition, it is also well established that oestrogens significantly promote breast carcinogenesis (Lipovka and Konhilas, 2016). For examples, mice treated with increased dosage of oestrogens have a higher risk of breast cancer in an earlier study (Russo and Russo, 1996). Similarly, in a recent study, using breast cancer cell lines, oestrogens appeared to be associated with enhanced cell growth and motility in oestrogens responsive MCF-7 cell line through an oestrogen-mediated integrin

intracellular pathway (Ho, Chang et al., 2016). Based on clinical studies, patients with oestrogen-positive tumours had a better survival when subjected to standard oestrogens-depletion therapy compared to patients with oestrogen-negative tumour, suggesting the association of oestrogen to the progression of this disease (Finn, Martin et al., 2016, Rose, Vtoraya et al., 2003). A majority of all breast cancers diagnosed are oestrogen receptor-positive (Zhang, Man et al., 2014). Oestrogens exert their activity by passing through the cell membrane, binding to its specific oestrogens receptor in the cell's nucleus, and transforming the receptor into an activated transcription factor. Activation of oestrogen-specific elements (EREs) which lie within the promoter region of specific regions occurs and this ultimately led to the regulation of various expressions of genes (Yaşar P, Ayaz G et al., 2016).

The biological effects of oestrogens action often resulted in rapid cell proliferation and differentiation. Cells exposed to excessive oestrogens grow in a rapid manner in which often the DNA damage left uncorrected, and these eventually lead to accumulation of DNA mutations (Tubbs and Nussenzweig, 2017, Yasuda, Sakakibara et al., 2017). Therefore, oestrogen-stimulated cells with mutations in critical genes often induce breast cancer. However, the majority of patients developed resistance to hormone oestrogen therapy (Patani and Martin, 2014). It has been shown that both oestrogens and some signalling growth factors involved in the oestrogen resistance breast cancer pathogenesis, specifically modulating the growth factor receptor expressions (Kovats, 2015, Mills, Rutkovsky et al., 2018). Numerous studies have been attempted to clarify the contributions of these growth factors to the disease progression.

Epidermal growth factors (EGFs), is one of the first growth factors being investigated, whereby overexpression leads to activation of a series of cascade reactions and formation of protein complexes that promotes breast cancer cell proliferation and migration (Lim, Li et al., 2016, Zhao, Ma et al., 2018). Deregulation of fibroblast growth factors (FGFs) expression on the other hand has been suggested to protect the mutated cells from apoptosis, potentially leading to survival of breast cancer cells and eventually cancer progression (Meng, Vander Ark et al., 2015, Wheler, Atkins et al., 2016). The insulin-like growth factors (IGFs) that are aberrantly expressed in malignant breast epithelial cells, have been found to interact synergistically with the oestrogen signalling network through the autocrine and paracrine interactions within the complex breast tumour microenvironment (Christopoulos, Msaouel et al., 2015, Scully, Firth et al., 2016). The development of specific biological inhibitors should provide new approaches in breast cancer management, yet it comes with numerous challenges due to the integrated signalling networks by the various growth factors.

2.1.3 Molecular Breast Cancer Subtypes

It is important to group breast cancers into their relevant subtypes as different subtypes require different therapeutic strategies (Yersal and Barutca, 2014). Classical methods of grouping breast cancer are mainly based on morphology characteristics performed by pathologists. Techniques based on immunohistochemistry markers correlated with clinic pathological data (size of tumour, grade of tumour, and involvement of lymph nodes) are often used to group the tumours into different categories (Eliyatkın, Yalçın et al., 2015).

However, this traditional way of classification of breast cancer has its own limitations. Over the years, the advancement of molecular platforms for gene expression profiling (e.g. microarrays) have successfully group the breast cancers not based on their physical attributes, but rather on the intrinsic molecular characteristics subtypes (Hassanpour and Dehghani, 2017). With this, it enables the heterogeneity of breast cancer even within a single tumour or tumour of the same type to be classified (Desai, Torous et al., 2018). Thus, in the new taxonomy, which was pioneered by Sorlie et al, 2000, the tumours were classified into four molecular subtypes (Table 2.1).

2.1.4 Risk factors, Diagnosis and Treatment

Breast cancer is a complex and multifactorial disease (Sabry, Mostafa et al., 2018). The development of breast cancer involves many stages and often over a duration of time. A majority of women feel an unusual lump in their breasts which is often painless as the first sign of breast cancer. Risk factors associated with high incident rates of breast cancer have been identified and these include age, dense breast tissue, physical inactivity and obesity, reproductive and menstrual history, history of menopausal hormone replacement therapy, alcohol consumption and smoking (Sun, Zhao et al., 2017). In addition, recent research has implicated those genetic alterations (inherited mutations of *BRCA* genes or other breast cancer susceptibility genes) increase the likelihood of women to develop breast cancer (Apostolou and Fostira, 2013, Castro, de Santiago et al., 2015). Still, many women with no apparent risk develop the disease. Therefore, having a healthy lifestyle may help to lower the risk of getting breast cancer.

Table 2.1. Summary of breast cancer molecular subtypes. (Eliyatkın, Yalçın et al., 2015)

Molecular subtype	Receptor expression	Histologic grade	Prognosis	Prevalence	Response to treatment
Luminal A	ER+, PR+, HER2-	Low (1 or 2)	Good	23.7%	Endocrine
Luminal B	ER+, PR+, HER2-	Intermediate (2 or 3)	Intermediate	38.8%	Trastuzumab
Her2 +	ER-, PR-, HER2+	Intermediate (2 or 3)	Poor	11.2%	Chemotherapy
Triple negative	ER-, PR-, HER2-	High (3)	Poor	12.3%	Chemotherapy

Current diagnostic approach involves thorough physical examination on the overall breast texture, size, skin condition, and any changes in the nipples (Wöckel, Albert et al., 2018). Further screening test involves the use of low dosage of X-ray known as mammography to examine the breasts for any cancer signs (Li, Zhang et al., 2016, Seely and Alhassan, 2018). Mammogram is normally useful for early detection of breast cancers in which treatment can be carried out most effectively. Nevertheless, this commonly used screening approach is hampered by insensitivity and misdiagnosis (Chan, Coopey et al., 2015). For instance, it has been shown that women with dense breasts are often misdiagnosed and it also cannot detect very small tumour in the breasts resulting in false negative result. Therefore, the use of mammogram for screening needs to be re-evaluated to minimise any unnecessary interventions resulting from the potential drawbacks of mammography. On the other hand, this problem is further compounded in developing countries because of diagnosis made in late stages in the majority of women, mammogram cost and limiting treatment availability (Sharp, Hippe et al., 2019). Currently, the best option for early intervention is prompt diagnosis. It is only in recent years that, molecular screening being introduced in breast cancer diagnosis as a result of advancement in molecular tools (Sokolenko and Imyanitov, 2018). Comprehensive molecular testing has been introduced due to the discovery of various ‘biomarkers’ that can be found only in tumour tissues (Kalia, 2015). The molecular markers for breast cancer include cancer-related oncogenes as well as tumour-suppressor genes, hormonal-based receptors, and protein biomarkers (Walsh, Nathanson et al., 2016). Commercially available tests such as MammaPrint Test and OncoTypeDx Test to analyse the activities of breast

cancer-related genes and their recurrence risks are one of the few examples of advanced molecular diagnostics tools using microarray technologies (Beumer, Witteveen et al., 2016, McVeigh, Hughes et al., 2014). The results from the tests are very beneficial and therefore enable better treatment decisions. Potentially novel microRNAs-based diagnostics have been recently incorporated as new molecular markers in predicting the behaviour of cancer (Nicolini, Ferrari et al., 2018, Tan, Liu et al., 2018). It is hoped that with improvements in microRNA expression profiling technologies, more new microRNAs as molecular markers could be discovered. These markers may be even more informative for prognostication of cancer and together with other molecular method possibly lead to the emergence of personalised medicine as what we know it now (Gao, Wang et al., 2018, Sun, Zhu et al., 2018).

Treatment recommendations for breast cancer depend on several factors which include the staging (tumour size and extent of spread), the molecular subtype of the tumour, age, and patient preferences (Waks and Winer, 2019). For early breast cancer stage, surgery to remove the tumour in one of the breasts or a bilateral mastectomy (removal of both breasts) is preferred (Garcia-Etienne, Tomatis et al., 2012). Surgery is also performed for those women in high risk group with *BRCA1* or *BRCA2* gene mutations as a form of preventive surgeries (Chiesa and Sacchini, 2016). For recurrent and metastatic cancer, adjuvant therapies which include chemotherapy, hormonal therapy, and targeted therapy are often followed surgery (Chan, Coopey et al., 2015). Chemotherapy normally acts by inhibiting the cell cycle of cancer cells, preventing them from growing (Punzi, Meliksetian et al., 2019).

Administrations of chemotherapeutic drugs post-surgery are performed to eliminate the remaining cancer cells (Karlsson, Cole et al., 2017). For early and locally advanced breast cancer, chemotherapeutic drugs used are tamoxifen, cyclophosphamide, docetaxel, and paclitaxel, in which combinations of 2 or 3 drugs are used for a more effective treatment (Anampa, Makower et al., 2015, Tong, Wu et al., 2018). For metastatic breast cancer, combination of chemotherapeutic drugs use includes cisplatin, doxorubicin, and 5-fluorouracil (5-FU) (Merino, Nguyen et al., 2016, Sledge, Neuberg et al., 2003).

In general, chemotherapeutic drugs work in all phases of the cell cycle act by either damaging the DNA of the cancer cells to prevent further cell division or interfering with the DNA and RNA activities. Hormonal therapy, on the other hand, is effective treatment for tumours (Luminal A and B) that are positive for oestrogen receptors (ER) and progesterone receptors (PR) for early or metastatic breast cancer (Rugo, Rumble et al., 2016). The therapy blocks the action of the hormones on breast cancer cells preventing recurrence in post-surgery treatment or causes death of the cancer cells resulting in shrinkage in tumour in the advanced stage (Tremont, Lu et al., 2017). Hormonal therapy is not effective for triple negative (oestrogen-receptor and progesterone-receptor negative) breast cancer (Al-Mahmood, Sapiezynski et al., 2018). Therefore, targeted therapies which target the specific genes or proteins of the cancer cells of triple negative breast cancer are used (Ahmed, Koval et al., 2019, Ju, Zhu et al., 2018). Some of the targeted therapies which have been proposed are VEGF inhibitors (Ribatti, Nico et al., 2016), PARP inhibitors (McCann and Hurvitz, 2018), and EGFR-targeted therapies (Nakai,

Hung et al., 2016). These approaches have potential for improving the treatment of patients with triple-negative breast cancer despite this research is still at its early phase.

New types of targeted therapies include the use of monoclonal antibodies to treat a variety of human cancers (Weiner, Surana et al., 2010). For example, the monoclonal antibody trastuzumab (Herceptin) bind to HER2 receptor (one type of growth factor receptors) on cancer cells, preventing the HER2 proteins from binding and stop their growth (Maximiano, Magalhães et al., 2016). On the other hand, DNA repair genes are genes responsible for correcting damaged DNA in the cells. When normal repair mechanisms fail, mutations in other genes tend to occur and this together could eventually lead to formation of a tumour (Romero-Laorden and Castro, 2017). Treatment targeting DNA repair pathways in cancer cells are still in its infancy as challenges in finding the exact targets for DNA repair inhibitors to induce tumour death (Gilmore, McCabe et al., 2019, Kelley, Logsdon et al., 2014). Therefore, development of biomarkers for cancers for targeted therapies proves to be better guidance for approaches. Breast cancer is still a disease that is difficult to treat despite advancement in treatment modalities because of its late detection, high rate of recurrence, and drug-resistance cell populations within the tumours. Early diagnosis could be the key for a better management of cancer as it has been shown that the 5-year survival rate for Stage I breast cancer is 100%, followed by Stage II breast cancer which is estimated to be 93%, for Stage III breast cancer is about 72% and it is only about 22% survival rate for Stage IV breast cancers (Hawkes, 2019). Therefore, newly identified

biomarkers such as miRNAs targets could possibly represent an entirely new diagnostic and prognostic tool for breast cancer (Han and Li, 2018, Nassar, Nasr et al., 2017).

Apart from genetic changes, lifestyle and environmental factors have been proposed to be the main culprit of cancer (Grosso, Bella et al., 2017). A number of studies have found that some cancers are preventable with major lifestyle and dietary changes. Tobacco smoke which contains at least 50 types of carcinogens has been accounted for about 25-30% mortality in cancer, with 87% death from lung cancer (Islami, Torre et al., 2015). High consumption of red meat among westerners contributes to a higher risk of deaths in colorectal cancer (Turner and Lloyd, 2017). Interestingly, laboratory evidence has shown the association of dietary intake and the effects on our genetic makeup. A cross-sectional study has found that people with diets high in vegetables, fruits and whole grains expressed genes favouring protective effect on cancer development and reducing risk of chronic diseases to the western diet (Guénard, Bouchard-Mercier et al., 2017). Nevertheless, the association of lifestyle and dietary patterns to the genetic abnormalities that lead to cancer would require further investigation. Alternatively, the newly targeted screening approach and tailored therapeutics treatment are beginning to provide us with the insights into the molecular pathogenesis of the cancers.

2.2 Cancer Stem Cells (CSCs)

2.2.1 Definition and Origin of CSCs

Tumours are a group of heterogeneous cells resulted from uncontrolled cell division. It is proposed that there are two theories on the existence of CSCs (O'Brien, Kreso et al., 2009). The conventional first hypothesis known as the stochastic model assumes that every cell within the tumour has equal chance to be a CSC (Odoux, Fohrer et al., 2008, Song, Wang et al., 2017). During cell division, random mutation causes a selective advantage in a population of cells and the mutated cells gain the ability to self-renew to drive the tumour progression. On the other hand, the second hypothesis proposes that cells in tumours are organised in a hierarchy way, in which CSCs are organised at the highest hierarchy, and the remaining differentiated cells lie at the base (Rich, 2016). This small population of cells at the apex are biologically different having the ability to self-propagate to continuously replenish the pool of CSCs via symmetric division. The two possible models for CSCs are illustrated in Figure 2.3. Despite the different views on the origin of CSCs, the formation of tumour is not only driven by this subpopulation of cells within the tumour, but also the tumour microenvironment (Roato and Ferracini, 2018). It refers to the immediate environment whereby the cancerous cells interact with the non-cancerous cells which include fibroblasts, endothelial cells and immune cells to support the development of the tumour (Ahmed, Escalona et al., 2018, Ingangi, Minopoli et al., 2019).

Different alternative terms have been used in literatures to describe these small subpopulations of cells and the terms include cancer initiating cells, cancer maintaining cells, tumour propagating cells, and cancer tumorigenic cells (Bomken, Fiser et al., 2010). These subpopulations of cells are proposed to have high self-renewal capability and behave like normal stem cells to a certain extent (Ayob and Ramasamy, 2018). The current consensus for this subpopulation of cells is known as cancer stem cells (CSCs). Despite intensive studies on the origin, cause, and treatment, there are still many questions unanswered about the molecular mechanisms that drive the malignancy and recurrence of the cancer (Taniguchi, Suzuki et al., 2019). Emerging evidences have shown that CSCs exist within the tumour responsible for tumour induction, development, metastasis and response to treatments (Chang, 2016, Lytle, Barber et al., 2018). The first evidence that proposed the existence of CSCs came from hematopoietic stem cells study from primary human acute myeloid leukaemia (AMLs) (Lapidot, Sirard et al., 1994). Interestingly, the fraction of a subpopulation of leukaemia cells isolated from transplanted experiment in immunocompromised mice bearing surface markers CD34+CD38- were highly tumorigenic compared to the CD34-CD38+ fractions. To distinguish such cells from its normal colony-forming cells, they were described as leukaemia-initiating cells or better known as leukemic stem cells (LSCs).

The gold standard to define CSCs is based on the ability recapitulate the parental tumours by performing the serially transplantable xenograft

experiments. Subsequent works performed by Bonnet D and Dick JE in 1997, again strengthen the findings where the stem-like cells in leukaemia were able to induce tumours in serially transplantable experiments while differentiated leukemic cells do not. Since then, CSCs have also been successfully described in solid tumours, with breast cancer being the first being reported (Al-Hajj, Wicha et al., 2003). The cells extracted from metastatic pleural effusions were fractionated into two populations based on CD24 and CD44 surface markers. Interestingly, fractions of cells that displayed CD24-CD44+ were reported to have significantly higher tumorigenic potential than CD24+CD44-/+ cell fractions when injected into NOD/SCID mice. Since then, several similar researchers have been conducted in other solid tumours including colon cancer, prostate cancer, hepatic cancer, pancreatic cancer and brain cancer (La Porta, 2012, Tirino, Desiderio et al., 2013). Ultimately, the procedure of purifying the CSCs from the solid tumours are the same using cell surface markers particular for the tumour type and the stem cell-like characteristics were assessed using mice xenograft transplantation experiments (Duan, Qiu et al., 2013, Kim, Pearson et al., 2016).

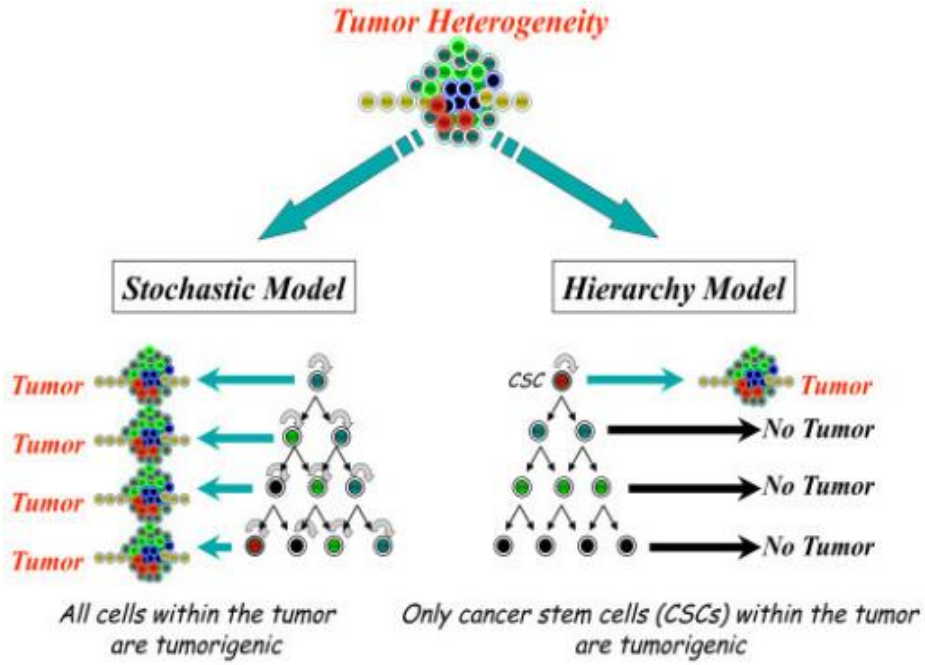


Figure 2.3. Stochastic vs hierarchical models of origin of cancer (Chandler and Lagasse, 2010). In the stochastic model, every cell in the tumour has the same potential to generate a tumour. In the hierarchy model, only a small subset of cells (cancer stem cells) in the tumour can generate a tumour.

2.2.2 Cancer Stem Cells Isolation

The existence of CSCs has been demonstrated in various tumours and cell lines. However, isolation of these subpopulation of cells is still difficult as they are typically exist less than between 0.1% and 5.0% of the total tumour cells (Hirata, Hatano et al., 2019, Huang and Rofstad, 2017). Standardised strategies to identify and isolate CSCs populations from tumours are not well established.

Isolating CSCs populations based on expression of cell surface markers are still the widely used methods in all the solid tumours (Kim and Ryu, 2017). Several surface markers have been well documented as potential CSCs markers, including CD34, CD24, CD44, CD133, and EpCAM (Kong, Lyu et al., 2018, Zhao, Li et al., 2017). All these markers are either used as single marker or in combination for a more effective identification. Many other CSCs surface markers have been identified, however, potential usage of those markers needs further investigations as some are controversial (Fanali, Lucchetti et al., 2014). The CD34, is a sialomucin-like adhesion transmembrane protein cell surface marker, was first used to detect hematopoietic progenitor cells as it is rarely expressed in normal tissues (Sidney, Branch et al., 2014). The use of CD34 in CSCs was first reported in human AML studies whereby leukemic stem cells were known to express CD34+CD38- capable of initiating tumour in immunocompromised mice (Lapidot, Sirard et al., 1994). In another study, CD34+ expression was used as predictive marker of relapse in oral squamous cell carcinoma (Kademani,

Lewis et al., 2009), suggesting of a population of primitive squamous stem cells. Over the time, CD34 used in different combinations like CD34+CD38-CD123+ have been found to be associated with higher chemoresistance with the upregulation of drug resistance proteins (Pearce, Taussig et al., 2005, Zhao, 2016). CD24, on the other hand, is a glycosylated 35-60 kDa sialoprotein which is usually expressed in human neuronal lineages (Pruszak, Ludwig et al., 2009). High expression of CD24 has been found in human embryonic stem cells suggested its use as a marker for CSCs (Shakiba, White et al., 2015). Since then, CD24 has been frequently used in combination with CD44 to identify breast CSCs (Horimoto, Arakawa et al., 2016, Li, Ma et al., 2017) and recently proposed as a new potential biomarker for human nasopharyngeal carcinoma (NPC). It was demonstrated that CD24+ NPC cells had enhanced invasion ability and as little as 500 cells of CD24+ inoculated in NOD/SCID mice were able to induce tumours (Yang, Wang et al., 2014). Increasing evidence implicated that CD44, a transmembrane glycoprotein which acts as a receptor for hyaluronan, plays critical role in CSCs in regulating the tumour microenvironment (Yan, Zuo et al., 2015). Beside the important use as biomarkers in breast cancer studies, CD44 is a valuable prognostic marker in various type of cancers due to its major role in self-regulating the properties of CSCs (Bartakova, Michalova et al., 2017, Wang, Tang et al., 2019). Till now, the sensitivity and specificity of identifying CSCs based on specific surface markers are being challenged as a single available specific marker is still unavailable (Pattabiraman and Weinberg, 2014, Zhao, Li et al., 2017). Owing to this inconsistency, reports have suggested using a combination of positive

and negative markers, peculiar for each tissue of origin would be a better approach in identifying CSCs.

The side population (SP) method based on the abilities of CSCs to exclude Hoechst 33342 dye was also previously reported (Hou, Liu et al., 2018, Yasuda, Torigoe et al., 2013). The rationale of this technique is based on the capability of the cells to extrude this dye via ATP binding cassette (ABC) transporters, which are also present in normal stem cells counterparts. Therefore, high effluxion SP cells are often reported to be rich in CSCs. However, the sorting of side population is no longer a preferred method for the isolation of CSCs as there were reports indicating the conflicting association between the SP fraction and CSCs (Broadley, Hunn et al., 2011, Burkert, Otto et al., 2008). For example, it has been reported that all the four gastrointestinal cancer cell lines examined, SP and non-SP populations were similar in their experiments and did not differ with respect to their stem-cell like characteristics (Burkert, Otto et al., 2008). Conversely, in another study, the SP cells demonstrated features similar to CSCs as well as higher resistant to chemotherapy (Naito, Wakabayashi et al., 2016, Yang, Fan et al., 2015). This contradicting result of SP further questions the association of SP and CSCs, which would require further investigations on the reliability of this technique.

Increased activity of aldehyde dehydrogenase enzyme (ALDH1) has been associated with CSCs (Mariel E. Toledo-Guzmán, Miguel Ibañez Hernández et al., 2019). ALDH is an important enzyme that is involved in

cellular mechanisms responsible for cellular detoxification, therefore protecting stem cells from oxidative stress substances (Vassalli, 2019). It was originally used for isolation and identification of haematopoietic stem cells (Ginestier, Hur et al., 2007). Subsequently, this enzyme has been successfully used to isolate CSCs from non-CSCs from a variety of cancers, including that from breast (Dehghan Harati, Rodemann et al., 2019), liver (Chang, Lo et al., 2019), colon (Vishnubalaji, Manikandan et al., 2018), lung (Shao, Sullivan et al., 2014) and pancreas (Ishiwata, Matsuda et al., 2018). Particularly interesting results were the comparative data done on ALDH1 positive and ALDH negative subpopulations within the whole tumour demonstrated that ALDH1 positive fraction were more tumorigenic and metastatic than the negative fraction (Li, Ma et al., 2017). However, there is a limited usage of ALDH1 as the functional marker as it is an intracellular metabolic marker that may question its specificity for targeting true stem cells in cancer.

2.2.3 CSCs and Cancer Recurrence

Researchers linked normal stem cells and CSCs together as they are able to divide indefinitely. These two stem cells are also able to maintain their stem cell pools by generating the bulk population of cells through differentiation and self-renewal (Hu, Mirshahidi et al., 2019). Though they shared similar features, CSCs are fundamentally different as they are not normal; in other words; they are responsible for cancer recurrence. Many recurrent cancer cases were due to different biology of the tumours which accounted for the most therapeutic failures and eventually death in patients.

The underlying molecular mechanisms related to chemotherapeutic resistance and cancer relapse remain a mystery. Based on recent observations, it has been hypothesised that the current chemotherapy or radiotherapy could have enriched the CSCs in the tumour while reduce the overall tumour bulk (Phi, Sari et al., 2018). This could be that these conventional therapies were not able to destroy the CSCs but instead aid in conversion of non-CSCs into CSCs within the bulk of tumour. For example, hepatocellular carcinoma cells exposed to carboplatin-mediated treatment were found to increase the stemness characteristics and gain in pluripotency genes (Hu, Ghisolfi et al., 2012). Similarly, in another study, breast cancer patients and mice models were found to have stem cell markers such as nuclear NOTCH and ALDH1 elevated post neoadjuvant therapies (Liu, Yang et al., 2018).

Apart from being chemoresistance, CSCs have been known to sustain less DNA damage in adverse tumour microenvironment (Schulz, Meyer et al., 2019). In normal circumstances, upon radiotherapy, DNA lesion causes cells to be arrested at specific cell cycle checkpoints to allow time for repair to take place. However, if the lesion is too significant, self-apoptotic pathways will be activated that eventually lead to apoptosis (Giglia-Mari, Zotter et al., 2011). During DNA damage, radiation-induced oxidative stress such as reactive oxygen species (ROS) is generated within the cells (Redza-Dutordoir and Averill-Bates, 2016). Remarkably, CSCs have underlying protective mechanism that can scavenge the ROS and enable repair of DNA in the cells. For example, CSCs can neutralise some ROS more efficiently than non-CSCs

thus sparing them from radiation effect (Ding, Li et al., 2015, Shi, Zhang et al., 2012). Therefore, having this more efficient DNA repair mechanisms, CSCs are able to evade chemotherapy or radiotherapy and increase their ability to metastasise (Mishra, 2018). Recent research looking at metabolic activity of cancer cells, found that CSCs subpopulations are driven by mitochondrial energy rather than glycolysis to fuel the high energy requirement of CSCs during the EMT (Fiorillo, Sotgia et al., 2019). This alteration to the metabolic requirements could suggest the use of novel metabolic targets to curb cancer recurrence and also provide insights to the complex nature of CSCs requiring further elucidation.

Therefore, eliminating CSCs could only be successful by inhibiting stemness properties of CSCs in order to limit recurrence. Among areas of interest could be looking at therapies targeting pathways of tumorigenesis such as EMT-related pathways, DNA repair mechanisms and drug-resistance pathways as they are among potentially targetable therapeutic pathways.

2.2.4 Spheroid as Enrichment Model for CSCs

The spheroid culture is a way of culturing cancer cells in a three-dimensional spatial conformation *in vitro* (Weiswald, Bellet et al., 2015). It involves culturing the cells in an anchorage independent manner which has high resemblance to the *in vivo* solid tumours (Sant and Johnston, 2017). The 3D spheroids share many features of the biological *in vivo* tumour

microenvironment in terms of their histopathological, functional and microenvironmental features than that of their monolayer counterparts (Kim JB, Stein R et al., 2004, Weiswald, Bellet et al., 2015). As most of the cells in a monolayer setting only allow the cells to grow as single sheet of cells typically in a culture dish, thus the results from monolayer cultures are often misinterpreted. Mounting evidences have demonstrated that the cells derived from monolayer cultures had limited cellular responses as they differ morphologically and physiologically from cells in 3D cultures (Edmondson, Broglie et al., 2014, Lv, Hu et al., 2017, Pickl and Ries, 2008).

The spheroid culture was first discovered in a neurosphere assay when the undifferentiated multipotent neural cells were successfully grown and maintained in a suspension culture (Reynolds and Weiss, 1996). Since then, it has been used in many researches specifically in the assessment of cancer drug resistance and discovery (Carver, Ming et al., 2014, De Angelis, Francescangeli et al., 2019). The CSCs hypothesis suggests that most of the stem cells have the ability to grow anchorage independently supplemented with the necessary growth factors on non-adherent surfaces (Mehta, Novak et al., 2018, Zhu, Wang et al., 2018). Interestingly, emerging evidences have demonstrated that spheroids grown in this way are enriched for CSCs (Arab-Bafrani, Shahbazi-Gahrouei et al., 2016, Bagheri, Memar et al., 2018, Yang, Zhang et al., 2018). Since then, this anchorage-independent sphere culture of stem cells has been used to enrich stem cells in human breast stem/progenitor cells known as mammospheres (Grimshaw, Cooper et al., 2008, Wang, Lv et

al., 2014), prostate cells known as protospheres (Bahmad, Cheaito et al., 2018, Collins, Berry et al., 2005), human hepatoma cell lines known as hepatospheres (Cao, Zhou et al., 2011), and human colon cancer known as colospheres (Shmelkov, Butler et al., 2008). The serum-free culture condition was thought to keep the CSCs in their undifferentiated stage facilitating their enrichment, as only stem cells were able to survive and proliferate in that condition, and non-stem cells may possibly undergo anoikis (Kruyt and Schuringa, 2010, Lin, Lee et al., 2012, Shaheen, Ahmed et al., 2016).

Anoikis is an intracellular mechanism occurred when a cell initiates cell death upon detachment from the extracellular matrices or neighbouring cells (Gilmore, 2005). The less oxygen availability within the core of spheroid architecture, a condition known as hypoxia is also known to modulate the cells to acquire stemness (Yun and Lin, 2014). Therefore, this serum-free enrichment technique provides a useful tool to select and expand tumourigenic CSCs and can also be used to generate sufficient amount of stem cells for functional studies such as drug response. Cells with higher self-renewal capability would be able to grow as floating spheres (Mehta, Novak et al., 2018). As a result, spheroid-enriched CSCs models have been proposed to be the next tool for analytical endpoints studies looking at the therapeutic testing targeting CSCs populations (De Angelis, Francescangeli et al., 2019, Huang and Rofstad, 2017). Nevertheless, these cell models have not been comprehensively characterised despite their extensive use.

There were some critiques on the use of serum-free spheroidal culture for CSCs enrichment. During the initial CSCs development, clonal expansion is the more appropriate way to expand CSCs *in vitro* as the cells are more likely to be a single clone of cells (van Niekerk, Davids et al., 2017). However, this view has been challenged as the clone of cells could be a result of cell aggregation due to the spontaneous locomotion of free-floating single cells and also the movement of plates during daily cell culture maintenance that possibly lead to cell aggregation (Ishiguro, Ohata et al., 2017). As such, it is generally accepted that clonal and aggregates of cells might coexist and resulted in a heterogeneous spheroid. Therefore, research on CSCs on various types of cancers including breast, prostate, colon and melanoma has relied on cell aggregation method (Cui, Hartanto et al., 2016, Tancioni, Miller et al., 2015, Vermeersch, Wang et al., 2015). In addition, there were reports on unsuccessful enrichment of CSCs in certain cell lines or primary cultures (Ramgolam, Lauriol et al., 2011, Wu, Haihong Zhang et al., 2014). The conflicting results could be due to lack of standardised protocols for serum-free enrichment technique for CSCs. It should also be noted that serum free can only enrich CSCs, whereas functional assays are able to fully characterise the CSCs population. The CSCs subpopulations could still be present but unfortunately were not able to be captured by the existing functional and biomarkers assays for the analysed cell lines or tumours. Therefore, the most relevant characterisation methods for CSCs depend on the study aim and also the intrinsic properties of the cell lines (Sancho, Barneda et al., 2016, Takeishi and Nakayama, 2016).

2.2.5 Breast Cancer Spheroid CSCs

Various methods have been reported to enrich CSCs from breast cancers using spheroidal cultures, but the core principle remains the same. Spheroidal cultures are based on anchorage independent methodology that minimises the interaction between the cells and the matrices (Antoni, Burckel et al., 2015). The aim of this is to prevent the cells from adhering to the underlying matrices, thereby allowing the cells to establish cell-cell adhesion resulted in spherical shape masses of cells. Different spheroidal enrichment of CSCs techniques have been previously reported including liquid overlay technique (Guo, Chen et al., 2019), gyratory rotation and spinner flask (Gencoglu, Barney et al., 2018), scaffold based cultures (Lee, Park et al., 2018), and spontaneous aggregation (Tevis, Colson et al., 2017). The liquid overlay technique, being the simplest and easiest way for tumour cells to achieve spontaneous aggregation to mimic solid breast carcinomas has been the most widely used technique (Gao, Wu et al., 2018). Accumulating evidences have demonstrated that spheroids generated using this technique maintain the biological properties of the intact tissues. Liang et al., (2018), demonstrated the use of MCF-7 tumour spheroids derived from liquid overlay technique were used for investigation of anti-CSCs nanoparticles targeted therapies. Wang et al. (2014) demonstrated successful enrichment of CSCs grown as spheroids under agarose non-adhesive culture system forming the basis for preclinical studies of cervical cancer. Similarly, Abboodi M. (2014) showed increased expression key breast CSCs markers in MDA-MB-231 spheroids when compared to that of the monolayer cells. Despite variability in the spheroid

generation technique, spheroids cultured in serum-free environment were said to be able to promote undifferentiated state of cancer cells.

2.3 MicroRNA (miRNA) Introduction

miRNAs are a group of short non-coding RNAs that are about 19 to 25 nucleotides in length (Acunzo, Romano et al., 2015). Being evolutionarily conserved within the mammalian genomes, they are found in all mammalian tissues responsible for various biological, developmental, pathological and oncogenic processes (Agarwal, Mulgund et al., 2015). Despite being small size, miRNAs is important in interfering gene expression at the post-transcriptional level, mainly by inducing mRNA degradation (O'Brien, Hayder et al., 2018). Each miRNA is known to regulate hundreds of target mRNA transcript. Therefore, the majority of the processes are miRNA-dependent regulation. Approximately, up to 2000 potential miRNAs have been identified in human according to miRBase version 22.1 (accessed June 2019), and they are predicted to modulate more than 60% of the human protein-output post transcriptionally. Recent reports have shown that miRNAs are not only present in cells and tissues but can be found in body fluids such as blood plasma and urine, collectively known as circulating miRNAs (Deb, Uddin et al., 2017, Wang, Luo et al., 2017).

2.3.1 miRNA Biogenesis and Role

The synthesis of miRNA biogenesis begins with the synthesis of single stranded RNAs decoded from the miRNA genes by RNA Polymerase II in the

nucleus of a cell. These single stranded RNAs or commonly known as primary transcripts miRNAs (pri-miRNAs) folded back into a hairpin stem-loop structure of about hundreds to thousands of nucleotides in length. Then, these structures are processed by protein complexes consist of Drosha and DiGeorge Syndrome Critical Region 8 (DGCR8) together with RNAse III enzyme that cleave the primary transcripts miRNAs to produce precursor miRNAs (pre-miRNAs). These 70-100 nucleotides long precursor miRNAs are then ready to be sailed to the cytoplasm via a protein transporter, exportin-5. To be part of the RNA interfering mechanism, the precursor miRNAs undergoes further processing by enzyme DICER that cleaves its hairpin stem-loop structure resulted in short double stranded miRNA: miRNA* duplex that later matures in cytoplasm. This mature duplex is then ready to be attached to a RNA interference silencing complex (RISC) that is associated with Argonaute and PAZ protein family members. The non-functional miRNA* is later presumably degraded by the cellular machinery in the cell (Lin and Gregory, 2015) (Figure 2.4).

The basic principle of mechanism of action of miRNAs is through complementarity sequence binding to the specific 3'-untranslated sequence sites on target mRNA. The silencing mechanisms which occur on mRNA post-transcriptionally involve either blocking its translation via deadenylation or decapping or through degradation of the mRNA targets (Di Leva and Croce, 2013b). The immediate effect is affecting the amount of protein encoded by the gene with resulting change in behaviour of the cell. Accumulating evidence has shown that miRNAs have a crucial function in cellular processes for

normal development, suggesting that aberrations of miRNAs could be involved in various human diseases, including cancer (Hayes, Peruzzi et al., 2014, Peng and Croce, 2016). Establishment of the expression profiles could help to reveal cluster of miRNAs that underexpressed or overexpressed among breast cancers as well as the different derivatives of cell types, thereby identifying miRNAs that could improve our understanding of breast cancer carcinogenesis.

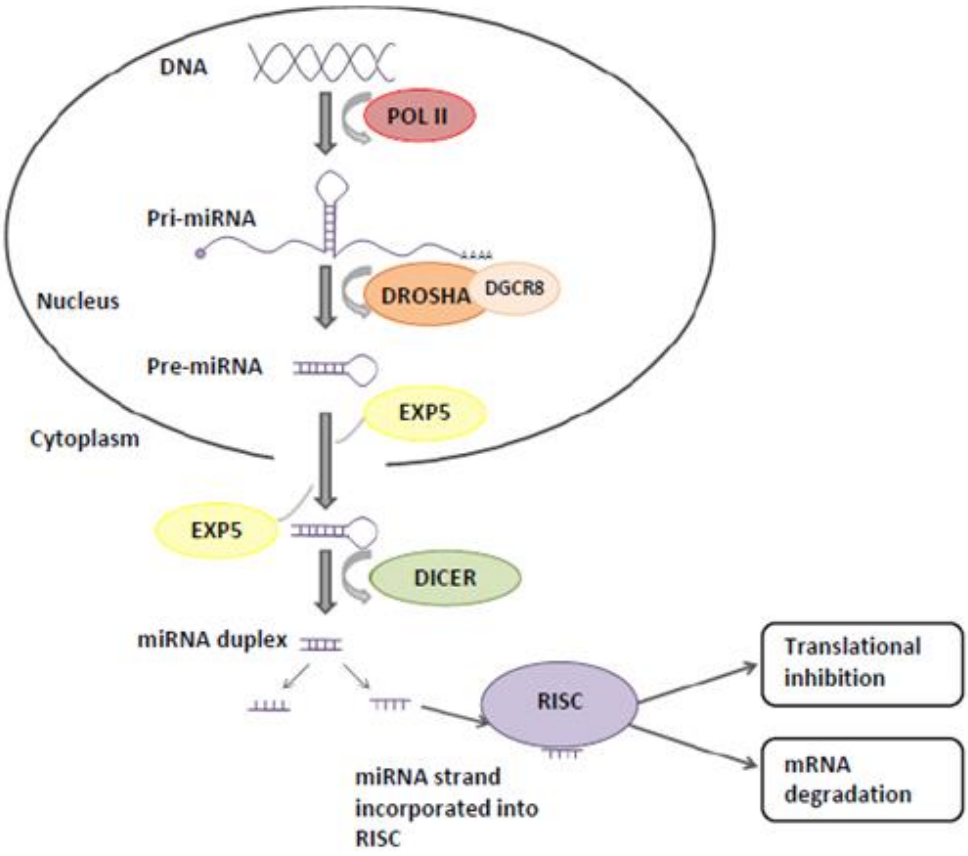


Figure 2.4: MicroRNA biogenesis pathway and miRNA function (Kristensen, 2015)

2.3.2 MiRNA Profiling Methods

The establishment of miRNA expression profiling has drawn much attention as miRNAs have significant roles as the main key regulators in gene expressions (Vishnubalaji, Hamam et al., 2015, Yan, Luo et al., 2016). Profiling of miRNAs is useful to identify miRNAs in the regulation of various biological processes, including establishment of miRNAs as important expression biomarkers (Gallo, Agnese et al., 2018, Panganiban, Lambert et al., 2019). Therefore, this has led to advanced development of miRNAs in the various fields of molecular diagnostics, including cancer. Various platforms have been developed for the profiling miRNA expression, and each of the method comes with its own advantages and limitations.

Currently, there are three commonly used approaches for miRNA expression profiling which include quantitative reverse transcription (qRT-PCR), miRNA microarray and miRNA-Next Generation sequencing (miRNA-NGS) (Pritchard, Cheng et al., 2012). An appealing aspect of using miRNA qRT-PCR approach is the established workflow protocol for existing real-time PCR users and also it allows absolute quantification as well as it is moderately highly specific and sensitive. The major drawbacks are the high variability of the optimal reaction conditions when subjected to high throughput capacity and it cannot be used to identify novel miRNAs (Pritchard, Cheng et al., 2012, Svoboda, 2015).

On the other hand, miRNA microarray is also an established and easily adapted method developed to complement the existing microarray workflow

(Wang, Howel et al., 2011). This hybridisation-based platform allows parallel analysis of large number of miRNAs in a single microarray chip. Individual miRNAs are detected based on complementary hybridisation of the fluorophores-labelled miRNAs with the short oligonucleotide probes arranged discretely as spots on the microarray slide. The advantages of miRNA microarray are that it is cost-saving compared to the other profiling methods and has a relatively time-saving protocol for microarray experienced users. However, current limitations include non-specific complementary binding, lower specificity to NGS, difficulty to be used in absolute quantification and less open source software tools for analysis (Koshiol, Wang et al., 2010, Love C. and S., 2013). A comparison of the different methods for miRNA profiling is in Table 2.2.

The latest advancement in next generation sequencing has allowed the development of miRNA-NGS resulted in precise identification of miRNA sequences that differ by single nucleotide (Buermans, Ariyurek et al., 2010, Tam, de Borja et al., 2014). This third approach which is known as miRNA-NGS begins with miRNA sample preparation from sample of interest, involves stringent quality control on the extracted miRNA, followed by establishment of miRNA cDNA libraries before subjecting them to sequencing workflow. A comprehensive bioinformatics analysis of the resulted sequence data is required to identify the sequence reads and functionally annotate the reads (Bisgin, Gong et al., 2018). Therefore, major limitations with this technology are the high cost, and the requirement of significant computational aid for data analysis, and the difficulty for absolute quantification (Hu Y., Lan W. et al.,

2017). Despite these known limitations, miRNA-NGS is still the preferred approach as it is highly sensitive and accurate for detection of known miRNAs as well as exploration of novel miRNAs. The NGS platform that is being applied in our study is the Ilumina platform (Liu, 2018, Luo, Tsementzi et al., 2012). This platform is one of the most reliable high throughout sequencing platforms in the market that offer massive parallel sequencing with high sensitivity and has the with widest dynamic range for miRNA profiling studies (Coenen-Stass, Magen et al., 2018, Lawless, Foroushani et al., 2013). In our work, the miRNA expression profiling enables genome-wide analysis of miRNA expression among samples of our interest. A clearer understanding of the mechanisms involved in miRNA deregulation in human cancer, specifically in breast will contribute greatly to the development of new cancer markers and miRNA-based strategies in cancer diagnosis, prognosis and treatment (Balatti, Tomasello et al., 2018).

Table 2.2: A comparison of different methods for miRNA profiling

(Catuogno, Esposito et al., 2011)

Detection method	Sensitivity	Specificity	Throughput	RNA input	Main feature
qRT-PCR	High	High	Low	<ng	Advantages • Comparatively inexpensive • Quantitative Limitations • Limited profiling • Labour intensive • Prone to contamination
miRNA microarray	Low	Low	High	ng-µg	Advantages • Comprehensive profiling • Time-saving protocol • Used in clinical diagnosis Limitations • Measure only miRNA relative abundance (not absolute quantification) • Less open software tools for analysis • Relatively expensive
Next generation sequencing	High	High	High	ng-µg	Advantages • Comprehensive profiling • Quantitative • Multiplexed Limitations • Requires significant computational aid for data analysis • Comparatively expensive • Highly labour intensive

2.3.3 MiRNA and Cancer

It is now well known that regulation of miRNAs is associated with every aspect of biological processes in the cells such as proliferation, apoptosis, differentiation, and tumorigenesis (Iorio, Ferracin et al., 2005). Therefore, deregulation of miRNAs is a common feature in human diseases, which include cancer, and these deregulations have been the main topic of cancer research (Paul, Chakraborty et al., 2018). The first complete evidence connecting miRNAs and cancer was identified by Calin et al., in 2004. It was observed that miR-15/ miR-16 clusters were the target genes of the 13q14 and deleted in 68% of chronic lymphocytic leukaemia patients. Following this initial finding, deregulation of miRNA expression has been reported in many instances in human cancer, specifically located in cancer-associated genomic regions (Esteller, 2011, Peng, Dai et al., 2013, Pereira, Marques et al., 2010).

Being the key players in numerous cellular processes, the patterns of regulation of miRNAs have direct impact on the production of tumour suppressive proteins and oncoproteins in cancer. Differential expressions of miRNAs in cancer disease suggest their complex roles as either oncogenic miRNAs (oncomiRs) or tumour-suppressive miRNAs (TS-miRs) (Vannini, Fanini et al., 2018). Overexpression of oncomiRs in cancers promotes cancer development by down-regulating TS-genes enables cancer cells to resist cell death and proliferate continuously (Frixa, Donzelli et al., 2015). On the other hand, underexpression of TS-miRs promotes tumorigenesis by activation of oncogenes enabling cancer cells to acquire proliferation and metastasis ability

(Vlachos, Zagganas et al., 2015) (Figure 2.5). Accumulating evidences have also implied that deregulation of miRNAs is involved in various signalling pathways as one miRNA can target more than one protein-coding gene (Treiber, Treiber et al., 2019, Voorhoeve, 2010). Thus, miRNA expression profiles have increasingly become important to explore the potential biomarkers for cancer diagnosis and prognosis and targeted therapy (Balacescu, Sur et al., 2018).

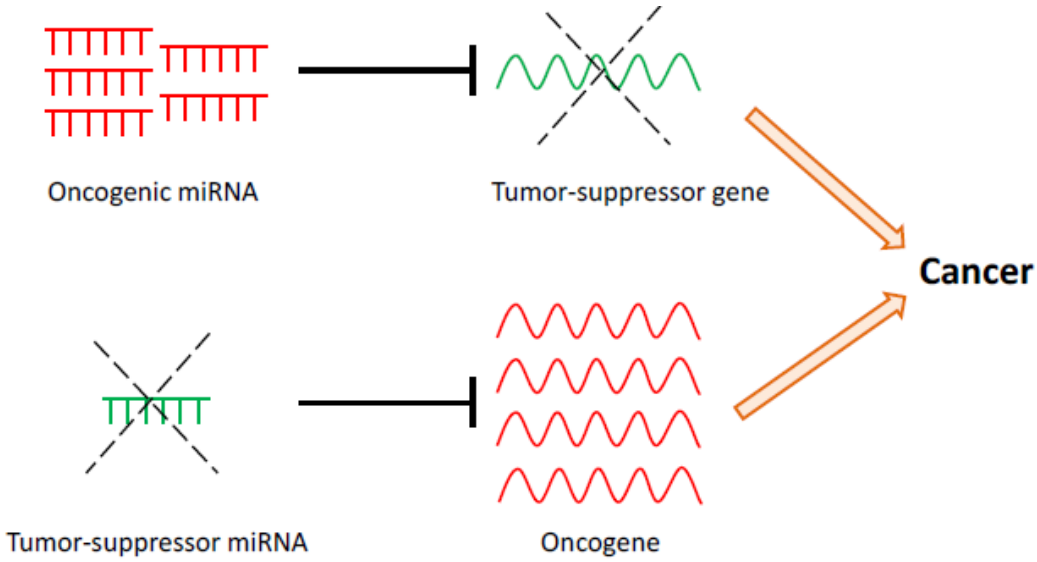


Figure 2.5. The two important roles of miRNAs either as oncogenic or tumour-suppressive roles (Kristensen, 2015)

2.3.4 miRNA and Breast Cancer

In breast cancer, aberrant expression of miRNAs has been investigated using various methods to identify miRNA signature characteristics that regulate breast cancer tumorigenesis in different breast cancer subtypes. One of the first miRNA expression profiling that successfully classified breast cancer subtypes was from analysis on 99 primary human breast tumours, 5 normal tissues and 33 breast cancer cell lines using microarray technology (Blenkiron, Goldstein et al., 2007). Of the 309 miRNAs analysed, 133 miRNAs whose expression was significantly deregulated were found to differentiate normal tissues from breast tumours. A large number of miRNAs were also found to be differentially expressed between the breast cancer subtypes and some miRNAs were found to be associated with clinic pathological features, which may serve as valuable prognostics biomarkers. In another study, miRNA microarray analysis in 76 breast cancers and 10 normal breast samples revealed specific clusters of deregulated miRNAs including miR-125b, miR-10b, miR-145, miR-21 and miR-155 that can correctly separate tumour from normal breast tissues (Iorio, Ferracin et al., 2005). Recent investigation on three specific oncoMiRs, miR-21, miR-222 and miR-372, conducted in 137 breast cancer patients, 60 benign lesion patients and 38 healthy controls using qPCR miRNA assays concluded that miR-372 had the highest diagnostic efficacy (Swellam, Zahran et al., 2019). Discovery of specific miRNAs that are co-expressed in body fluids and in the breast cancer tissues would benefit miRNA-based non-invasive early detection.

One of the major reasons for breast cancer complexity is due to their multiple subtypes characterised by distinct morphological, pathological, and clinical features (Dai, Li et al., 2015). With advancement in molecular profiling, it is also possible to link certain miRNA expression profiles with the breast cancer subtypes. In a large scale study focused on triple negative breast cancer (TNBC) or basal subtype, miR-21, miR-155, miR-200a family and miR-210 were the most significantly up-regulated miRNAs, whereas let-7b, let-7c, miR-126, miR-145 and miR-205 were the most significantly down-regulated revealed from 165 primary TNBC and 59 normal tissues (Cascione, Gasparini et al., 2013). In addition, validated studies have also been carried based on the TNBC miRNA expression profiles that demonstrated their reliability as prognostic biomarkers and therapeutic targets. High expression of miR-210 was found in TNBC as well as in the tumour microenvironment compared to luminal subtype, and its upregulation was associated with poor prognosis (Bar, Merhi et al., 2017, Lü, Mao et al., 2017). On the other hand, meta-analysis of miRNA microarray revealed nine prominent miRNAs associated with breast tumours, and three of these miRNAs (miR-130a, miR-92a-, miR-211, and miR-500a) were significantly overexpressed in luminal subtypes (Oztemur Islakoglu, Noyan et al., 2018). MiRNA profiling is also able to predict specific miRNAs associated with the ER, PR and HER2 status within the luminal subtypes. The miRNAs which were ER specific include miR-342, miR-299, miR-217, miR-190, miR-135b, PR specific include miR-218, miR-520g, miR-377, miR-527-518a, miR-520f-520c and HER2 specific were linked to miR-520d, miR-181c, miR-302c, miR-376b, miR-30e (Lowery, Miller et al., 2009). Of all this, miR-342 was of particular interest as further

analysis revealed that this miRNA was able to distinctively differentiate between luminal B and TNBC subtypes. Importantly, miRNAs signatures obtained from expression profiles does not only contribute to subtype classification, but also reveal their potential as prognostic and diagnostic tools. The ability to determine the molecular breast cancer subtypes is important as it facilitate the planning of effective targeted therapies.

Breast cancer metastasis occurs when the cancer spreads and invades other parts of the body from its site of origin (Jin and Mu, 2015). Metastatic breast cancer contributes to the high mortality rate of all breast cancer-related deaths (Narod and Sopik, 2018). Several studies have reported the potential use of miRNAs as biomarkers along the process of metastasis in breast cancer. In particular, miRNAs that promote metastasis include miRNA-23b/27b/24 cluster (Ell, Qiu et al., 2014), miR-182 (Yu, Lei et al., 2016), miR-103/107 (Martello, Rosato et al., 2010), miR-9 (Chen, Price et al., 2013), miR-132 (Anand, Majeti et al., 2010), miR-210 (Andorfer, Necela et al., 2011) , and miR-374a (Cai, Guan et al., 2013). For examples, pre-clinical animal model study demonstrated that the miRNA-23b/27b/24 cluster contributed to increased cell migration and invasion in breast cancer lung metastasis by targeting PSAP (Ell, Qiu et al., 2014). Overexpression of miR-182 was found to enhance invasion leading to bone metastasis by targeting SMAD7, activating TGFß-induced EMT pathway, suggesting the potential use as miRNA-based cancer intervention (Yu, Lei et al., 2016). MiRNA identified as miR-9 was found to target metastasis suppressor E-cadherin, and thus promote tumour invasiveness and metastasis in breast cancer carcinoma (Chen, Price et al.,

2013). Importantly, miRNAs contribute to the metastasis process and they are envisioned as the next anti-metastasis targeted therapies for metastatic cancer (Yoo, Kavishwar et al., 2017).

2.3.5 MiRNA and Breast CSCs

It is well known that breast cancer relapse is associated with a group of subpopulations of cells known as CSCs within breast tumour which have the ability to self-renew the same way as normal stem cells (Turdo, Veschi et al., 2019). Considering the importance of CSCs, characteristics of miRNAs expression signature on breast CSCs could possibly help to predict the molecular regulation of breast CSCs. Based on previous studies, there were some shared features of breast CSCs and normal mammary stem cells irrespective of the breast tumour subtypes (Liu, Cong et al., 2014). Importantly, the shared miRNAs were involved in regulating the stem cell functions in both normal and breast CSCs such as capacity for self-renewal, differentiation into non-self-renewing cells and the up-regulation of specific stem cell genes (Takahashi, Miyazaki et al., 2014). For example, miR-200 family mediates self-renewal ability of stem cells through interactions with ZEB-transcription factors and found to be downregulated in normal and breast cancer stem cells (Feng, Wang et al., 2014). Members of the miR-200 family were also found to target p53 which then increase the expression of EMT and stem cell characteristics (Kim, Veronese et al., 2011). Using ALDH+ and CD44+ to isolate the stem cell population from MCF-7 breast cancer cell line, Han et al., 2012, determined that miR-21 was highly expressed in CSCs-like cells by promoting EMT through suppressing HIF-1α. MiR-600, on the other hand,

was recently found to act as a bimodal switch in regulating breast CSCs property by be either promoting self-renewal or differentiation through Wnt signalling (El Helou, Wicinski et al., 2014). In addition, miR-181 serves as a critical regulator in breast CSCs survival as overexpression of miR-181a was reported to induce sphere formation via the ataxia telangiectasia mutated (ATM) gene (Wang, Yu et al., 2011). Recently, the expression of miR-181a has been found in serum of breast cancer patients, proposing the use of this miRNA as non-invasive diagnosis biomarker, though this application is still in its early infancy (Fan, Chen et al., 2017, Yang, Tabatabaei et al., 2017).

To sum up, miRNAs have pivotal roles in regulating breast CSCs including self-renewal, metastasis, invasion and drug resistance via several core CSCs regulatory pathways. Till now, several of these regulatory pathways have been identified that are important to maintain CSCs within the tumours, including Wnt, Hedgehog, and Notch signalling pathways (Imamura, Mukohara et al., 2015). Wnt signalling enables some cancer cells to acquire extended self-renewal capacity as the normal stem cells. One well reported study found that miR-142, which was highly expressed in human breast CSCs compared to the normal murine mammary tissues, activates Wnt signalling pathway accompanied with mammary dysplasia in the *in vivo* models (Isobe, Hisamori et al., 2014). Aberrations in the Wnt pathway also enhance self-renewal activity which has been reported in other types of cancer CSCs such as leukemic CSCs, hepatocellular CSCs and colorectal CSCs, possibly regulated by other miRNAs (Dahal Lamichane, Jung et al., 2019, Katoh, 2017, Pandit, Li et al., 2018). Extensive investigations have revealed alterations in Hedgehog

signalling pathway have led to activation of stemness-determining genes such as Nanog, Oct 4 and Sox 2 reported in the maintenance of self-renewability in solid tumours CSCs (Cochrane, Szczepny et al., 2015, Koury, Zhong et al., 2017). In a recent *in vivo* study, Hui et al., (2018), performed a randomised trial on therapeutic targeting on Hedgehog pathway on TNBC tumour stroma. Their work indicated that inhibiting the pathway lead to remodelling of the tumour stroma which in turn significantly reduced the CSCs proportion and improved the mice survivability. Successful inhibition of Hedgehog pathways for targeted CSCs therapy has also been reported in pancreatic and brain CSCs (Valentina, Mattia et al., 2016, Wang, Ma et al., 2016).

Notch pathway is one of the most important developmental pathways in stem cell growth and differentiation (Kopan, 2012). Activation of this pathway in CSCs has caused unrestricted cell proliferation, and avoidance of apoptosis in cancer cells, which therefore maintain the survival of CSCs (Abel, Kim et al., 2014). Epithelial-to-mesenchymal transition (EMT) is believed to be the main mechanism to be regulated by these three core CSCs pathways (Cai, Cao et al., 2018). This mechanism is an evolutionary conserved program across many species of mammals during normal embryonic development that allows epithelial cells to undergo morphological changes (Brabletz, Kalluri et al., 2018). EMT promotes cancer cells to acquire stem cell characteristics by breaking down the cell-cell and cell-extracellular matrix connections, allowing cancer cells to metastasise. Mani et al., 2008 first reported the connection between EMT and CSCs by upregulating stem cell markers in breast cancer. Over the years, induction of EMT not only promotes metastasis of cancer but

also contributes to drug resistance which makes tumours more difficult to cure (Shibue and Weinberg, 2017, Wang, Wei et al., 2016). Along with transcription factors, miRNAs play a key role in regulating cells that have undergone the EMT process (Czerwinska and Kaminska, 2015). A multi-study integration analysis based on data mining in breast CSCs highlighted eight miRNAs including miR-200c, miR-21, miR-204, miR-30c, miR-34a, miR-10b, miR-520c, and miR-373 were tied to EMT (Rahimi, Sharifi-Zarchi et al., 2019). Of particular interest is miR-204, which was predicted to be the miRNA that determines the breast CSCs fate as it was consistently upregulated in all types of mammospheres investigated and targeted EMT transcription factors (Rahimi, Sharifi-Zarchi et al., 2019). Since EMT plays a major role in sustaining CSCs, finding the potential regulatory miRNAs responsible in EMT is particularly important to relate this mechanism to the core CSCs regulated pathways.

At present, deregulation of these pathways and their associated miRNAs provide novel and safer targeted therapeutics approaches in both CSCs and bulk tumour cells. Yet, other CSCs pathways especially those associated with breast cancer subtypes remain to be explored. It is important that future investigations apply high-throughput sequencing technologies that would be able to provide valuable insights regarding new CSCs-related molecular alterations to target CSCs. Consequently, these findings could contribute to the development of innovative personalised breast cancer treatment in combination with conventional strategies.

2.4 Study Rationale

Generally, CSCs have been known as the subpopulation of cells that have the ability to self-renew and give rise to differentiated cells that constitute the heterogeneity of the tumour tissue. These subpopulations of cells are the key drivers for tumour progression, tumour recurrence and possess high resistance towards most of the common chemotherapeutic drugs. Much effort has been made to isolate or enrich CSCs from a wide range of human tumours, but not much was addressed on the comprehensive phenotypic and molecular characteristics of the CSCs derived from two most established breast cancer cell lines, MCF-7 and MDA-MB-231. Therefore, in this work, we tried to prove if serum-free spheroid culturing technique allows the generation of CSCs and if it does, are there differences between the two breast cancer subtypes. MiRNA transcriptomic profiling was performed to further infer the potential regulation and contribution of miRNAs toward the CSC phenotypic features.

2.4.1 Objectives of this Study

In the present work, we tried to prove if anchorage-independent cell culture techniques allow the generation of spheroid cultures and if these cultures were enriched for cells with phenotypic properties of CSCs. Although, deregulation of miRNAs has been reported to be aberrantly expressed in breast cancer, specific miRNAs involved in spheroid-enriched CSCs have not been well defined. Therefore, in pursuit of this goal, the specific objectives of this study are:

1. To enrich CSC-like cells through spheroidal and serum-free culture from breast cancer cell lines, MCF-7 (luminal subtype) and MDA-MB-231 (basal subtype).

2. To characterise the enriched cells based on phenotypes of CSCs.

3. To establish the miRNA expression profile of enriched spheroidal breast cancer cells, and to identify the differentially expressed miRNA groups, relative to respective parental breast cancer cell lines.

4. To predict the functional annotations of the differentially expressed miRNAs groups based on bioinformatics analyses

CHAPTER 3

MATERIALS AND METHODS

3.1 Flow Chart of Methodology

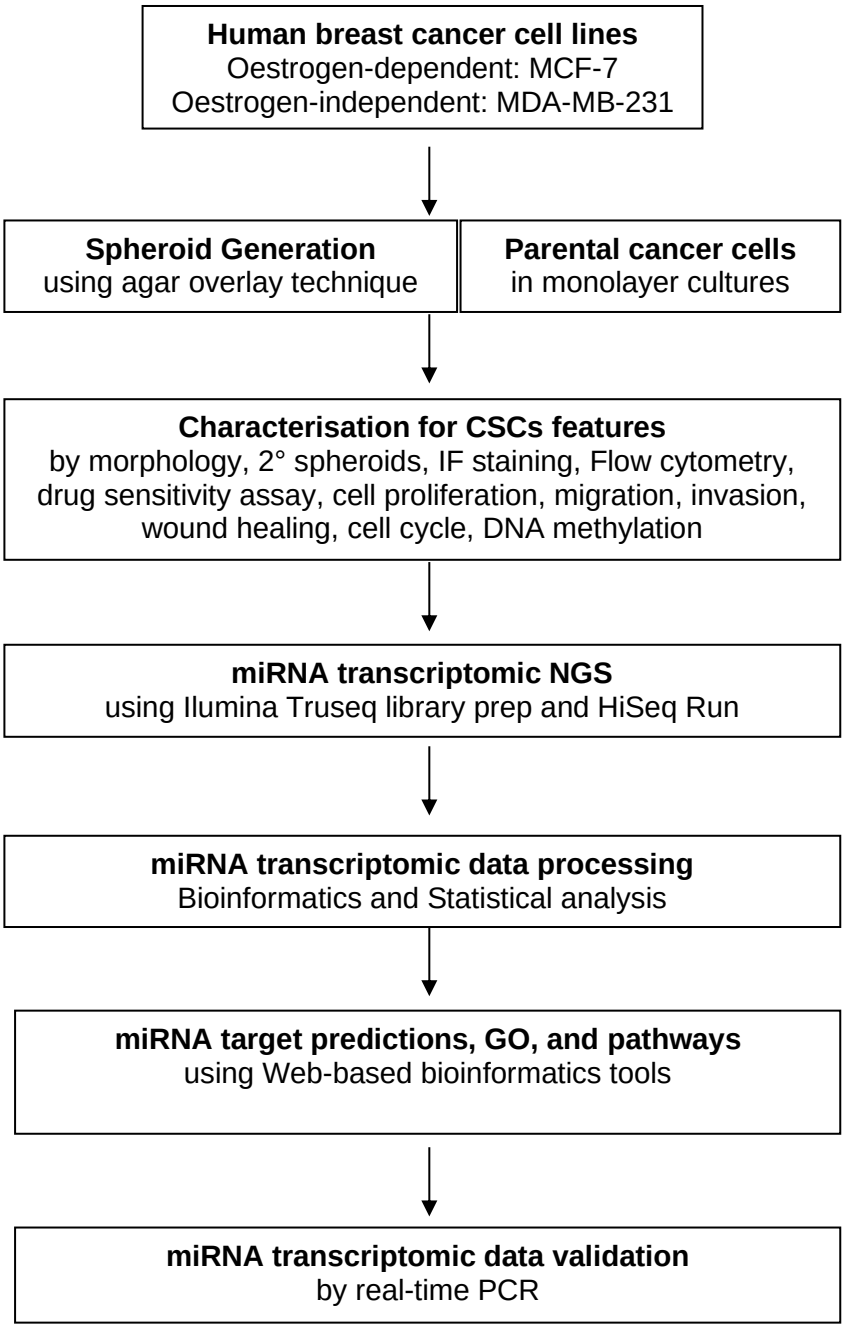


Figure 3.1. Flow chart of research methodology of the study.

3.2 Cell Culture

General guidelines for culturing of cell lines were undertaken in a Class II Biological Safety Cabinet (ESCO, Singapore). Aseptic techniques were performed to reduce the probability of contamination from the environment to the sterile work area. The work surface was wiped with 70% ethanol before and after work. Ultraviolet light was used to sterilize the exposed work surfaces in the cell culture hood between uses. Cell culture incubator (ESCO, Singapore) set at 37 °C and supplied with 5% CO_2 was used for the maintenance of the cultured cell lines. All the cell observation was recorded and captured using an inverted phase-contrast light microscope (Nikon, Tokyo, Japan). Cell counting was performed using an automated cell counter Countess II (Invitrogen, USA). All culture bottles and tips were sterilised in an autoclave machine prior use in culture work. Some media, supplements and reagents were filter-sterilised to prevent microbial growth in the cell culture.

3.2.1 Breast Cancer Cell Lines

Two human breast adenocarcinoma cell lines, MCF-7 and MDA-MB-231 were purchased from American Type Culture Collection (ATCC, Washington, USA) for the use in this study. MCF-7 (Catalog number: HTB-22) is an oestrogen-dependent human breast cancer cell line whereas MDA-MB-231 (Catalog number: HTB-26) is an oestrogen-independent human breast cancer cell line.

3.2.2 Preparation of Culture Medium

The Roswell Park Memorial Institute (RPMI-1640) cell culture media were prepared in accordance to the product sheet information. RPMI-1640 (Sigma, Missouri, USA) purchased in powder form was first dissolved in 90% of the final required volume in double distilled water. Sodium bicarbonate (Sigma, USA), weighed 2.0 g was then added for each litre of final volume of medium being prepared. The pH of the medium was then adjusted to pH 7.2 to 7.4 using 1N of HCl, before brought the solution to final volume. The medium was then filtered sterilised using a membrane filter unit with a porosity of 0.22 µm (TPP, Trasadingen, Switzerland) attached to a vacuum pump system. To make a complete growth medium, the filtered-sterilised medium was supplemented with 10% (v/v) heat-inactivated foetal serum (FBS) (Gibco, Gaithersburg, USA), 1% antibiotic containing 100 I.U/mL penicillin and 100 ng/mL streptomycin (GIBCO, USA), and 2mM L-glutamine (Biowest, Nuaillé, France).

3.2.3 Thawing Frozen Cells

The breast cancer cell lines were stored in cryovials and shipped to our laboratory in frozen form. Upon receiving, the cryovials were placed in liquid nitrogen vapour phase at a temperature below -130 °C. To initiate the culture, a frozen cryovial was removed from liquid nitrogen and thawed rapidly by gentle agitation in a 37 °C water bath for approximately 2 minutes. The vial was then removed from the water bath and decontaminate by spraying with 70% ethanol

before opening it. The thawed cells were then diluted slowly with 10 mL pre-warmed growth medium. The cell suspension was then centrifuged at 2000 rpm for 10 minutes to remove the cryoprotectant agent dimethyl sulfoxide (DMSO) which is present in the shipping medium. The supernatant was then discarded, and the cell pellet was resuspended in 2 mL fresh pre-warmed complete growth medium before transferred into a 75 cm^2 tissue culture flask containing 10 mL complete medium. The culture flask was then incubated at 37 °C under a 5% humidified CO_2 atmosphere.

3.2.4 Maintenance of Culture and Subculturing

Both MCF-7 and MDA-MB-231 breast cancer cell lines were maintained in RPMI-based complete medium. For maintenance purpose, the cells were examined under phase-contrast light microscope daily for any microbial contamination and to check the attachment of the cells to the bottom of the flasks. Culture medium was changed twice a week, 15 mL for each flask. When the cells reached 70% confluency, the cells were ready for subcultured. The medium was discarded from the culture flask. The cells were then washed with 2 mL 1X PBS by gently rocked the culture flask forth and back for several times. The wash solution was then removed and discarded. Two mL of pre-warmed dissociation reagent TrypLE (GIBCO, USA) was then added to the side of the flask, and gently rocked the flask to get a complete coverage of the cell layer. The culture flask was then incubated at room temperature for 2 to 3 minutes and the cell detachment was confirmed using the microscope. When 90% of the cells had detached, 2 volumes of pre-warmed complete medium

were added to the detached cell suspension to deactivate the trypsin. The detached cell suspension was then transferred to a 15-mL conical tube and centrifuged at 2000 rpm for 10 minutes to pellet down the cell suspension. The cell pellet was then resuspended in 1 mL pre-warmed complete growth medium and a sample of 20 µL of cells was removed for counting purpose. Total number of cells, cell concentration and percent viability were then determined using the Countess Automated Cell Counter. A subcultivation ratio of 1:3 and 1:6 are recommended for MCF-7 and MDA-MB-231 breast cancer cell lines respectively.

3.2.5 Population Doubling Time for Breast Cancer Cell Lines

To calculate the population doubling time (PDT), which is the time required for a culture to double in number, the breast cancer cell lines were first seeded in 6-well plates at 5000 cells per well. For ten consecutive days, one well was counted per each day, and the cells counted using Trypan blue assay. The experiment was carried out for 10 days. The growth curve of the cells was then plotted with cells per cm^2 against days after culture. The PDT was calculated during the exponential phase of the growth curve with the following formula:

$$PDT = T \ln 2 / \ln(\frac{Xe}{Xb})$$

Where,
T is the incubation time from the growth curve
Xb is the cell number at the beginning of the incubation time
Xe is the cell number at the end of the incubation time

3.2.6 Cryopreservation of Cultured Cells

Cell cryopreservation was carried out as an important backup supply for replenishing cells due to contamination and also to ensure a homogenous culture supply. Cells were detached according to the procedure used for subculturing (Yeap, S.K., 2010). The number of cells and percent viability were determined and only cells with viability rate >90% and in log phase were cryopreserved. Cells at concentration of 1 x 10^6 cells per mL were pelleted at 2000 rpm for 10 minutes. Supernatant was discarded and the cell pellet was resuspended in 1 mL freezing medium containing 10% (v/v) DMSO (MpBio, Irvine, USA) and 90% heat-inactivated foetal serum. The cell suspension mixed with freezing medium was then aliquoted into cryogenic storage vials (Greiner Bio-One, Rainbach im Mühlkreis, Austria). Cell freezing was carried out by keeping the cryovials in CoolCell cell freezing container (BioCision, New York, USA) in a -80 °C freezer for 4 hours before being transferred to liquid nitrogen for long term storage.

3.2.7 Mycoplasma Testing in Cell Culture

3.2.7.1 Sample Collection and Preparation of DNA

Mycoplasma testing based on polymerase chain reaction (PCR) was performed to detect if the cultures were contaminated with any mycoplasma contamination according to the method published (Drexler, H. G. and Uphoff, C. C., 2002). The cell suspension from the continuous culture was collected, washed in 1X PBS, and subjected to DNA extraction using GenEx DNA

Extraction kit (GeneAll, Seoul, Korea). Briefly, 1 x 10^6 cells of the cell line to be tested were harvested into a 1.5 mL microcentrifuge tube by centrifugation at 14,000g for 10 seconds. The supernatant was discarded and replaced with 300 µL of lysis buffer (Buffer AL), and incubated at 37 °C water bath for approximately 1 hour or longer till no visible cell clumps in the tube. One µL of RNase solution at concentration of 20 mg per mL was then added to the lysate and the sample was inverted for 5 times before incubated for another 5 minutes at 37 °C. Protein precipitation buffer (Buffer PP) was then added to the mixture and the tube was vortexed vigorously for 20 seconds before chilled on ice for another 5 minutes. The tube was centrifuged at 14,000g for 1 minute and the supernatant was carefully transferred to another 1.5 mL microcentrifuge tube containing 300 µL of isopropanol. The suspension was gently mixed by inversion until the white thread-like strands of DNA formed. The DNA was then finally concentrated using 70% ethanol, air-dried before dehydrated using 30 µL of 10 mM Tris-Cl, pH 8.5. The concentration of the extracted DNA was determined by measuring the absorbance value at 280 nm using a Nanophotometer (Implen, München, Germany). The purity was estimated by the ratio of the absorbance value at 260nm divided by the reading at 280 nm. Only DNA with a purity of A260/A280 of 1.8 to 2.0 was chosen for the subsequent work. The integrity of the DNA was also examined by 1% agarose gel electrophoresis. Five µL of 50ng/µL of the extracted DNA was mixed with 1 µL 6X loading gel (Thermo Fisher Scientific, Waltham, USA) and subjected to 1% agarose gel electrophoresis at 80V for 40 minutes in 1X TAE running buffer. GeneRuler DNA 1kb ladder (Thermo Fisher Scientific, USA) with known lengths was run in another well of the same gel. The gel was

then stained with 3X Gelred nucleic acid (Biotium, Fremont, USA) for 15 minutes before visualised with UV Bioimaging system (UVP, Upland, USA). Samples with DNA bands at above 10 kb were considered unsheared (intact) and were used in these experiments. The good integrity DNA was stored at -20 °C prior to the mycoplasma test.

3.2.7.2 PCR Reaction

Primers specific for mycoplasma PCR were obtained from First Base Laboratories, Serdang, Malaysia. The primers sequences as listed in Appendix A. Forward primers at 5 µM each (Myco-5') and reverse primers at 5 µM each (Myco-3') were then mixed in distilled water. Internal control DNA extracted from mycoplasma-positive samples was kindly provided by Dr.Yeap Swee Keong, Insitute of Bioscience, Serdang, Universiti Putra Malaysia. Good integrity DNA that was previously extracted was diluted at 50ng/µL. The MyTaq Red Mix ready-to-use PCR mastermix (Bioline, UK) containing all the necessary reagents was used. The PCR reaction is summarised in Table 3.1. The reaction tubes were briefly centrifuged and incubated at 95 °C for 3 minutes, followed by 35 cycles of 95 °C for 20 seconds, 55 °C for 20 seconds, 72 °C for 45 seconds, and finally 72 °C for 5 minutes. The reaction was performed using a Veriti 96-Well Fast Thermal Cycler (Thermo Fisher Scientific, USA). A negative control was included in each run. The PCR products were then subjected to a 1.0% agarose gel electrophoresis at 80 V for 40 minutes before the gel was stained and visualized with UV Bioimaging

system (UVP, USA). A band at 502-520 bp indicated mycoplasma-positive samples.

Table 3.1. Mycoplasma PCR reaction mixtures

Reagent	Volume (µL)	Final concentration
MyTaq Red Mix (2X)	10.0	1X
Nuclease-free water	7.0	-
Forward primer mix (5 µM)	0.5	0.125 µM
Reverse primer mix (5 µM)	0.5	0.125 µM
Template DNA (50 ng/ µL)	2.0	100 ng
Total Volume per Sample	20.0	

3.3 Generation of Spheroid Cultures

3.3.1 Preparation of Plates Coated with Agar

Agar-coated plates were prepared a day before the initiation of spheroid cultures. Agar Gelzan powder (Sigma, USA) was prepared at 1.0% (w/v), completely dissolved in 50 mL distilled water. The diluted agar solution was then autoclaved at 121 °C for 20 minutes using liquid phase cycle. The sterilised agar solution was then warmed at 60 °C with occasional stirring. Then, the agar solution was brought to culture hood and pre-warmed Dulbecco's Modified Eagle's F-12 Ham (DMEM-F12) (GIBCO, USA) culture medium was added at 1:1 ratio. Using a multichannel pipettor, 100 µL of the agar solution was then dropped into the wells of a 96-well flat bottom tissue culture plate. This action of pipetting was repeated until all the wells were covered with the agar solution. The agar solution produces a semi-clear gel and the coated plates were air-dried in the culture hood for approximately 15 minutes to ensure complete gelation. The coated plates were then stored in 4 °C before use.

3.3.2 Spheroid Generation

MCF-7 and MDA-MB-231 spheroids were generated based on the standard overlay tumour spheroid cultivation in agar-coated 96 tissue culture plates to prevent cell attachment. Cells were detached according to the procedure described in the subculturing section. Cells were seeded into each well at concentration of 5 $x10^4$ cells/mL in serum-free medium. Cell

aggregation was carried out by centrifuging the plate at 2000 rpm for 5 minutes and the plates were incubated at 37 °C under a 5% humidified CO_2 atmosphere for four days. When the spheroids reached day 4, the spheroids were ready to be used for the subsequent downstream experiments.

3.4 Characterization for Cancer Stem Cells Properties

3.4.1 Microscopic and Scanning Electron Imaging

Monolayer cultures of the parental MCF-7 and MDA-MB-231 breast cancer cell lines were examined daily to observe their morphology, the colour of the medium and the confluence state of the cells. Spheroid cultures of both breast cancer cell lines were also examined daily to assess the spheroid formation. All the cell observation was recorded and captured using a phase-contrast light microscope (Nikon, Japan) at 4X and 10X magnifications. For the analysis of the ultrastructure of the spheroids, scanning electron microscope (SEM) analysis was performed. Firstly, the spheroids were harvested from the 96-well agar-coated plates, washed with 1X PBS, before fixed in 2.5% (v/v) glutaraldehyde in 0.1M cocodylate buffer and kept in 4 °C. The samples were then post-fixed with 2% osmium tetroxide for 2 hours at 4°C, dehydrated through a graded ethanol series (35%, 50%, 70%, 80%, 85% and 95%) for 15 minutes each and twice at 100%. This dehydration process was then followed by immersing in pure acetone for 15 minutes each. Each step of dehydration was accompanied by a short centrifugation at 3000 rpm for 5 minutes to minimise the loss of samples during transferring. After further drying in fume

hood for 15 minutes, the specimens were then carefully mounted onto an aluminium stub using doublestick carbon tape. Specimens were then critically point dried (LEICA EM CPD030, Wetzlar, Germany) and sputtered gold coated (BAL-TEC SCD005, Wetzlar, Germany) before examined under SEM (JSM-6400; JEOL, Tokyo, Japan). The sample processing and viewing were carried in the Microscopic Unit, Institute of Bioscience, Universiti Putra Malaysia.

3.4.2 Secondary Spheroids Formation Assay

The resulted spheroids were collected and pooled by gentle centrifugation at 1800 rpm for 5 minutes. The supernatant was carefully removed using a Pasteur pipette, and the spheroids were then washed twice with 1X PBS. Spheroids were disaggregated into single cells suspension using Accutase (GIBCO, USA) incubated at 37 °C waterbath for 20 minutes. The cell suspension was resuspended by pipetting up and down with a 1 mL pipette to facilitate the disaggregation process. The reaction was stopped by adding complete culture medium and the cell suspension was passed through a 70 µm cell strainer (BD Falcon, New Jersey, USA) attached to a 50-ml centrifuge tube. The filtered cell suspension was then pelleted down and resuspended in 1 mL serum-free medium prior to cell counting. To generate secondary spheroids, single cells were plated at 200 cells per well in a 96-well Ultra-low attachment plate (Corning, New York, USA) in 100 µL serum-free growth medium. Fifty µL of medium was added every three days and the number of spheres for each well was evaluated after 14 days. The sphere formation

efficiency was counted in percentage as the number of spheres formed divided by the cells seeding density per well from first through the third generation. To further test on the *in vitro* self-renewal capability of the secondary spheroids formed, a single cell limiting dilution assay was performed with one cell per well placed to a 96-well ultra-low attachment plate (Corning, USA). The wells with one cell were observed and images of the propagation of the single cell forming a sphere from day 1 to day 14 were captured using a light phase-contrast microscope (Nikon, Japan). All experiments were performed in triplicates.

3.4.3 Immunofluorescence Staining

Parental and spheroid cells were prepared differently for the immunofluorescence staining. Parental cells were cultured as monolayers on Nunc Lab-Tek II chamber slides (Thermo Fisher Scientific, USA) until 80% confluency prior to fixation. For spheroids culture, the spheroids were collected in a microcentrifuge tube by centrifugation at 1800 rpm for 5 minutes prior to fixation. The samples were fixed in 4% cold paraformaladehyde in 1X PBS for 10 minutes at room temperature. Later the samples were washed three times in ice cold 1X PBS. For intracellular markers staining, the samples were permeabilised with 1X Permeabilization buffer (BD Biosciences, USA) for 15 minutes and then washed three times in 1X PBS for 5 minutes each. Sample blocking was then performed by incubating the cells with 1% bovine serum albumin in 1X PBS (Merck, New Jersey, USA) for 1 hour in room temperature. After washing three times in 1X PBS for 5 minutes each, the samples were

then incubated in the diluted antibody in 1X PBS in a humidified chamber for 2 hours at 4 °C. The primary mouse monoclonal antibodies anti-human conjugated to fluorescent dyes, consisted of CD24-PE, CD44-FITC, CD49f-FITC, Sox2-PE, Nanog-AF, and ALDH1-FITC. All the antibodies were purchased from Miltenyl Biotech, Woking, UK. The solution was then decanted and the cells washed three times in 1X PBS, 5 minutes each wash. The samples were then counterstained with 0.5 µg/mL DAPI (GIBO, USA) for 2 minutes before rinsed with 1X PBS. The slides were then blot dried and mounted with a drop of anti-fade mounting medium (Abcam, Cambridge, UK), covered with coverslips. The slides were then stored in the dark at 4 °C before captured using a fluorescent microscope (Carl Zeiss, USA).

3.4.4 Flow Cytometric CD44/CD24 and ALDH Activity Analysis

To determine the CSCs subpopulations in spheroids and parental, the expression of the surface markers CD44+/CD24- using anti-CD44-Fluorescein (FITC), and anti-CD24-phycoerythrin (PE) were measured. The monoclonal primary antibodies were directly conjugated antibodies purchased from Miltenyl Biotech, UK. One million cells were harvested using Accutase, gently disaggregated to a single cell suspension by pipetting up and down, passed through a 70 µm filter, and stained according to the manufacturer's protocol. Briefly, the cells were stained with the antibodies for 15 minutes at 4°C at 1:11 dilution and washed with ice-cold 1X PBS. Enzyme Aldehyde dehydrogenase (ALDH) activity was analysed using the ALDEFLUOR kit (Stem Cell Technologies, Vancouver, Canada) according to their manufacturer's protocol.

The dry ALDEFLOUR reagent was provided in a stable, inactive form, and DMSO was added to activate it. The disaggregated single cells were suspended with activated ALDEFLOUR reagent, one tube without diethylaminobenzaldehyde (DEAB) specific inhibitor and another one with the inhibitor to act as a control, both incubated at 37 °C for 45 minutes. Following incubation, the tubes were centrifuged to remove the supernatant and later resuspended in 0.5 mL ALDEFLOUR buffer on ice. For data acquisition, for surface markers CD44-FITC/CD24-PE and ALDH activity, flow cytometry was performed with a FACS Calibur Flow Cytometer (BD Biosciences, USA) and figures were analysed using Cell Quest Pro software.

3.4.5 Drug Sensitivity Assay

The sensitivity of parental and spheroids to chemotherapeutic drugs was measured by MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide) assay (Sigma-Aldrich, USA). Briefly, the parental and spheroids were seeded in 96-well microplates (TPP, Switzerland) on the same growth medium. Cytotoxicity testing on spheroids (3D format) were carried out based on published method (Ho, Yeap et al., 2012) while following standard methods for dissociated spheroids (2D format) and monolayer cells (Mosmann, 1983). The cell number to be used for monolayer cells was first determined from the cell titration experiment. Briefly, for spheroids in 3D format, the spheroids were transferred to a new 96-well plate and centrifuged at 2000 rpm for 5 minutes. For the spheroids in 2D format and the monolayer cells, the cells were dissociated into single cells, filtered, cell count and seeded

at 5 x10^3 cells/well in a 96-well plate a day before the drug treatments. Different concentrations of tamoxifen, cisplatin and doxorubicin (ranging from 0 µg/mL to 30 µg/mL for monolayer and spheroids (2D format) and ranging from 0 µg/mL to 100 µg/mL for spheroids (3D format) and diluents alone (the controls) were added to the cell cultures in a total volume of 200 µL/well. The chemotherapeutic drugs were purchased from Merck, USA. After 96 hours of incubation, the assay was performed by adding 20 µL of MTT solution (5 mg/mL in PBS) to each well, incubating 4 hours, adding 170 µL of DMSO to dissolve the formazan crystals, and reading the absorbance of each well at 570 nm in a microplate reader (Tecan, Männedorf, Switzerland). The morphology of the treated cells in the different conditions was captured at IC_{50} drug concentrations at 96 hours post-treatment using light phase-contrast microscope (Nikon, Japan). Untreated cells were cultured in parallel in the experiment. The cytotoxicity was determined by comparing the resulting absorbance with the mean absorbance of the control wells (considered as 100% viability) and was expressed as percentage of cell viability. The average values from triplicate readings subtracted the average values for the blanks were determined. The drug concentrations that caused inhibition of 50% cell viability (IC_{50}) were determined from the dose response curves. Experiments were carried out in triplicates.

3.4.6 Cell Proliferation Assay

Parental cells at 80% confluence and spheroids at day 4 were picked and dissociated into single cells using Accutase (GIBCO, USA). The

dissociated single cells were then cultured at 200 cells/well in 24-well ultra-low attachment plates (Corning, USA) in serum-free media. Half of the medium was changed every three days and cell proliferation was detected by alamarBlue Cell Viability Assay Reagent (Thermo Scientific, USA) at specific time intervals over the time course of three weeks. The length of incubation time and plating density were determined before the induction of the assay. Twenty µL of alamarBlue reagent was added directly to the cells in culture medium and the plates were incubated for 4 hours at 37 °C in a cell culture incubator, protected from direct light. After incubation, 100 µL of the culture medium from each well was aliquoted into a new 96-well tissue culture plate (TPP, Switzerland). No-cell control samples were run parallel in the experiment. The absorbance of each well was then read at 570 nm, using 600 nm as a reference wavelength in a microplate reader (Tecan, Switzerland). The proliferation assays were performed in triplicates. The cell proliferation rates for the dissociated cells originated from parental and spheroids were then calculated based on the percentage reduction of alamarBlue reagent using absorbance readings. The percentage difference in reduction between sample and control cells in the proliferation assay was determined by the formula as follows:

$$\frac{(\varepsilon ox)\lambda 2A\lambda 1 - (\varepsilon ox)\lambda 1A\lambda 2}{(\varepsilon RED)\lambda 1A'\lambda 2 - (\varepsilon RED)\lambda 2A'\lambda 1} \times 100\%$$

Where

εox = molar extinction coefficient of alamarBlue oxidized form (BLUE)
εRED = molar extinction coefficient of alamarBlue reduced form (RED)
A = absorbance of test wells
A' = absorbance of negative control well
λ1 = 570 nm
λ2 = 600 nm

3.4.7 Wound Healing Assay

Parental and spheroids were grown to confluence in a 6-well tissue culture plate (TPP, Switzerland) prior to the initiation of the assay. A line was drawn with a marker pen on the bottom of the plate. Wounds were initiated using a sterile 200 µL pipet tip through the cells moving perpendicular to the line drawn in the step before. The cells were carefully and gently rinsed with 1X PBS as not to lift off the sheets of cells. Images of the wounded were captured using an inverted microscope (Nikon, Japan) at 10X magnification just above and just below each line at 6, 12 and 24 hours post-wound initiation. The areas of the wounds at different time points were analysed using Image J software to calculate the wound healing rate.

3.4.8 Interpretation Results using Image J software

The images of the wounds taken at different time points of the different samples were first saved as .tiff files in a directory which is accessible in Image J software. Using the software, the .tiff file was loaded, edges were sharpened, and threshold were adjusted to black and white. Ensuring that the difference between the cells and the wound were still visible, the black and white image was then inverted, and the image was further analysed using "analysis particle size" (Figure 3.2). The percentage of the wounded area of the image was then determined by dividing the wounded area by the total area and multiply by 100%.

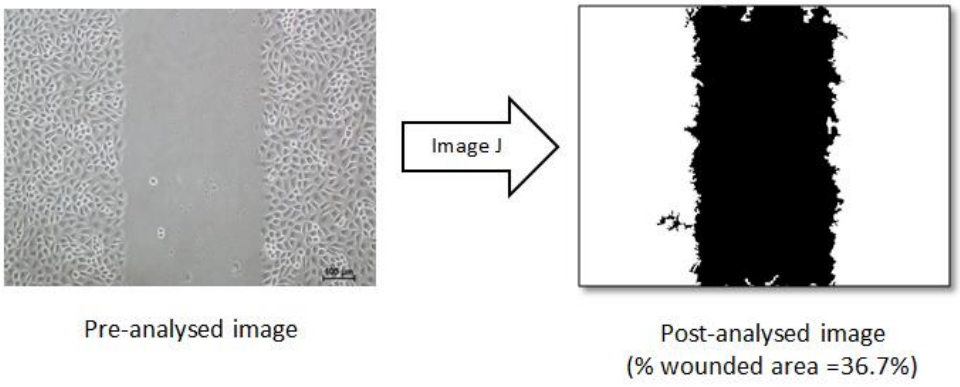


Figure 3.2. **Wound healing assay images analysed using Image J software**.

3.4.9 Tumour Invasion Assay

Parental and spheroids were grown to 80% confluence. The cells were starved in serum-free medium the day before the assay. The next day, the cells were dissociated into single cells with Accutase (GIBCO, USA), counted and added at 1 x 10^5 cells/well onto the top chambers of transwell inserts of 8-µm pore size filter (BD, USA) coated with basement membrane Matrigel (BD, USA) in a 24-well tissue culture plate. In the bottom of the chambers, DMEM containing 10% of FBS was added, and the cells were then cultured for 72 hours under standard culture conditions. At the end of the assay, the cells on the top surface of the filter were removed by using a cotton-swab, and the cells on the bottom of the filter were then fixed with methanol and stained with 0.5% (w/v) crystal violet (Sigma, USA) for 30 minutes. The chambers were then washed in 1X PBS, counted using an inverted microscope with either a 4X or a 10X objective lens and plotted as the percentage of invading of the total number of plated cells. The cell counting was performed using Image J software. For quantifying the invasive ability of the cells per filter as an absorbance, the dye was then extracted using 30% acetic acid and the concentration was measured at 590 nm in a microplate reader (Tecan, Switzerland). The percentage of invaded cells was measured using absorbance of the sample/absorbance of the control and expressed as percentage. The experiments were performed in triplicates.

3.4.10 Transmembrane Migration Assay

In the cell transmembrane migration assay, all the steps were carried out similarly to those in the invasion assay (Section 3.4.8) except for the Matrigel coating.

3.4.11 Cell Cycle Analysis

For cell cycle distribution analysis, parental and spheroids were performed using CycleTEST PLUS DNA Reagent Kit (BD Bioscience, USA) according to their manufacturer's instructions. Briefly, the cells were harvested and disaggregated into single cell suspension using Accutase (GIBCO, USA), filtered through a 70µm nylon mesh (BD Biosciences, USA), and counted to the concentration of 1.0 x 10^5 cells/mL. The cell suspension was then placed into a 17 x 100-mm tube, centrifuged for 5 minutes at 1800 rpm at room temperature. The supernatant was then aspirated, and the cell pellet was resuspended in 5 mL buffer solution by gently vertexing at low speed. The cells were then further washed with decreasing volume of buffer solution and then finally resuspended in 70 % ethanol by adding ethanol dropwise while vortexing. The cells were then kept at -20 °C for 30 minutes or until analysis. For frozen samples, the samples were thawed rapidly in water bath at 37 °C before subjected to staining. The resulting pellets were rehydrated with the buffer solutions, treated with 250 µL trypsin buffer (Solution A), treated with 200 µL trypsin inhibitor and RNAse buffer (Solution B) and stained with 200 µL propidium iodide stain solution (Solution C). The stained cells were then

filtered through a 50-µm nylon mesh into another new tube. The stained samples were then ready to be subjected for flow cytometric analysis using FACS Calibur Flow Cytometer (BD Biosciences, USA) and data were analysed using Cell Quest Pro software.

3.4.12 Global DNA Methylation Assay

Global DNA methylation levels were assessed by Methylflash Methylated DNA Quantification kit (Epigentek, New York, Sweden). MethylFlash which is colorimetric ELISA-based assay was employed in order to quantitatively measure genomic methylation of the parental and spheroids. Briefly, the DNA was extracted according to the protocol described in Section 3.2.7.1 and diluted to 100 ng in nuclease-free water. To facilitate the DNA binding onto the strip wells, 80 µL of Binding Solution (Solution ME2) was added to each well, followed by 100 ng of the DNA. DNA was bounded to the strip wells that are specifically treated to have a high DNA affinity. The solution was mixed by gently tilting from side to side to ensure the wells were evenly coated. The solutions were then incubated at 37 °C for 90 minutes. Solution ME2 was then removed and washed with 150 µL diluted wash buffer (Solution ME1) three times. Capture antibodies specific to the methylated fraction of DNA was then prepared and then added into the wells. The wells were then read using a microplate reader (Tecan, Switzerland) at 450nm. Negative control consisted of unmethylated polynucleotide containing 50% of cytosine and positive control consisted of methylated polynucleotide containing 50% of 5-methylcytosine provided in the kit were run parallel in the

experiment. Experiments were performed in triplicates. The amount of methylated DNA is proportional to the absorbance intensity measured based on the given formulas for relative methylation status of two different DNA samples. The formula for the relative quantification is as follows:

$$5-\mathrm{mC}\% = \frac{(\text{Sample OD} - \text{ME3 OD}) \div \mathrm{S}}{(\text{ME4 OD} - \text{ME3 OD}) \times 2 *\div \mathrm{P}} \times 100\%$$

Where

S = amount of input sample DNA in ng

P = amount of input positive control in ng

2* is a factor to normalize 5-mC in the positive control to 100%, as the positive control contains only 50% of 5-mC

3.5 miRNA Transcriptomic Profiling

3.5.1 Isolation of Total RNA Containing miRNAs

Total RNA with retention of miRNAs was extracted from the parental and spheroid cells using miRNeasy Mini Kit (Qiagen, Germantown, USA) in accordance to the manufacturer's instructions. The cells were trypsinized, washed with 1X PBS for complete removal of any residual medium, adjusted to 5 x 10^6cells per tube and stored in RNA later (Thermo Scientific, USA) at -80 °C. When the samples were ready for extraction, the samples were thawed quickly on ice, added with 2 mL 1X PBS, and the cell pellet was collected by centrifugation at 5000 rpm for 10 minutes. The cells were lysed in 700 µL

Qiazol Lysis solution by vortexing, and the cell lysate was then loaded onto a QIAshredder spin column placed in a 2 mL collection tube. The lysate was homogenized by centrifugation at 14, 000 rpm for 2 minutes before they were placed at room temperature for 5 minutes. Then, 140 µL chloroform was then added to the homogenate and the tube was mixed vigorously for 15 seconds and left at room temperature for 3 minutes. The sample was then separated into 3 phases by centrifugation at 12,000 rpm for 15 minutes. The 3 phases were consisted of; upper, colourless aqueous phase containing RNA; a white cloudy interphase containing DNA; and lastly a red, organic phase containing protein. The upper aqueous phase was then transferred to a new collection 1.5 mL tube, added with 535 µL of 100% ethanol and mixed thoroughly by pipetting up and down several times. Then, 700 µL of the sample was transferred into a RNeasy Mini spin column and was spun down at 12,000 rpm for 15 seconds. The flow-through was then discarded, and the tube was then spun again to ensure no ethanol carry-over. On-column DNA digestion was then performed by washing the membrane of the spin column once with 350 µL Buffer RWT. Reconstituted DNase I solution of 80 µL was then added directly onto the membrane of the spin column and the tube was placed on the benchtop in room temperature for 15 minutes. The spin column was then washed once with 350 µL Buffer RWT and twice with 500 µL Buffer RPE by centrifugation at 12,000 rpm for 15 seconds. The spin column was then transferred to a new 1.5 mL collection tube and the RNA was finally eluted with 30 µL RNase-free water by centrifugation at 12,000 rpm for 1 minute. RNA was then aliquoted into smaller volumes and stored at -80 °C for future use.

3.5.2 RNA Quantification and Integrity Check

The concentration and yield of the extracted RNA was determined by the absorbance value at 260 nm in a Nanophotometer (Implen, Germany). The purity of the extracted RNA was estimated by the ratio of the absorbance at 260 nm and 280 nm (A_{260}/A_{280}). When analysing total RNA with miRNA samples, the constant value of 33 instead of the value of 40 was manually entered into the nanophotometer. The presence of protein or phenol products produced high value at 280 nm, resulted in a lower A_{260}/A_{280} ratio. Therefore, RNA with values A_{260}/A_{280} ratios of 1.8 to 2.2 were deemed indicative of pure RNA and further quantified using Qubit RNA assay kit (Molecular Probes, Oregon, USA). In this assay, the extracted RNA was first diluted to 5 ng in 2µL nuclease-free water before mixed with 198 µL of Qubit working solution. The required low and high standard tubes were also prepared at 10 µL each and mixed with 190 µL of Qubit working solution. The final volume in each tube was adjusted to 200 µL. All the tubes were then mixed by vortexing 2 to 3 seconds and incubated at room temperature for 2 minutes. The measurement of the standards and samples were then performed on Qubit2.0 Fluorometer (Thermo Fisher Scientific, USA). The concentration of the original sample was calculated based on value given by the Qubit 2.0 Fluorometer and multiply by the dilution factor.

The integrity of the extracted RNA samples was also determined by Agilent 2100 Bioanalyzer using RNA 6000 Pico Kit (Agilent Technologies, Santa Clara, USA). The system uses a miniaturized version of acrylamide

gels in a microfluidics electrophoresis chip to separate RNA for analysis. The RNA 6000 Pico kit was chosen for its high resolution in the less than 100 nucleotide range that allowed the verification of miRNA retrieval in the total RNA samples. The Agilent 2100 Bioanalyzer was performed according to the guidelines. Briefly, the electrodes on the instrument were cleaned with RNase-free water for 5 minutes prior to use. To prepare the gel-dye matrix, the gel and the dye concentrate were allowed to equilibrate at room temperature for 30 minutes. The gel was centrifuged at 1500 g for 10 minutes while the dye concentrate was vortexed 10 seconds. In a new tube, 1 µL of the dye was mixed with the 65 µL of the filtered gel. The gel-dye mix was then spun at 13, 000 g for 10 minutes. The RNA samples were diluted to 0.5 ng in 1 µL nuclease-free water, heat denatured and then loaded into the designated wells of the pre-coated with gel-dye matrix RNA chip. Heat denatured ladder was also loaded at 1 µL in the designated well. The chip which can accommodate up to 11 samples were then vortexed for 1 minute at 2400 rpm and then subjected to run on the Agilent 2100 Bioanalyzer. The sample wells were then detected by fluorescence and electropherogram were created by the data analysis software for assessing RNA quality. RNA integrity number (RIN) scale ranged from 0 to 10, with 10 indicating maximum RNA integrity was also calculated by the algorithm of the instrument software. High, visible and distinctive peaks of 28S and 18S ribosomal RNAs demonstrated on the electropherogram indicated RNA samples that were intact (no sign of degradation). Only RNA samples with sufficient concentration (600 ng/µL) and RNA integrity >8 were used for next generation sequencing miRNA library preparation and subsequent data validation by real time quantitative

PCR. The threshold of RIN >8 was applied to ensure that only RNA of high integrity was used in this study. RNA samples were stored in small aliquots at -80 °C to minimize RNA degradation due to repeated freeze-thaw processes.

3.5.3 Multiplex miRNA Library Construction for Illumina Sequencing

The miRNA libraries were prepared using TruSeq Small RNA Sample Preparation Kit (Ilumina, San Diego, USA). TruSeq kit allow to use of 24 different index tags to make use of the Illumina multiplexing capability for analysis of the miRNA samples. The Illumina multiplexing system uses six-base indices to distinguish different samples from one another in a single lane of flow cell. The protocols for the miRNA library preparation make use of the common structure of the miRNA molecules. The adapters provided in the kit were directly ligate to the 5'-phosphate and the 3'-hydroxyl group of the mature miRNAs. After ligation, the libraries were then subjected to single stranded cDNA synthesis. Then, the cDNA was then PCR amplified using a common primer and a primer containing one of the 24 index sequences. All the generated miRNA libraries were then pooled and gel purified to generate a final library product before send for next generation sequencing. Good molecular biology practices were adhered throughout the preparation of RNA libraries for sequencing. All incubations were conducted in a C1000 Touch Thermal Cycler (Bio-Rad Laboratories, Hercules, USA). The protocol is summarised in the section 3.5.3.1.

3.5.3.1 miRNA Sample Prep Workflow

The figure 3.3 illustrated the preparation of one sample using Truseq Small RNA Sample Prep protocol.

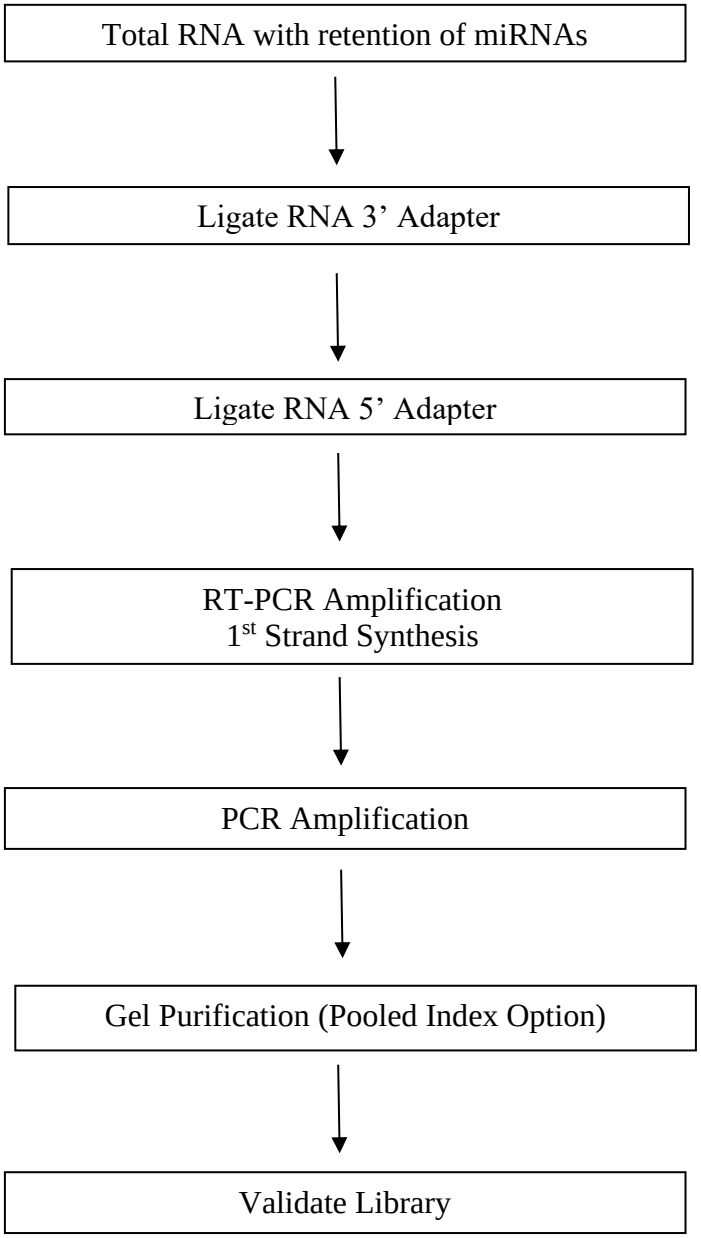


Figure 3.3. MiRNA library sample preparation using Truseq Illumina Kit.

3.5.3.2 Ligation of 3’ adapter

The protocol suggested using 1 µg of total RNA for the library preparation. However, after some optimization, 3 µg of total RNA was determined as the total RNA input. Good-quality total RNA was diluted to 600 ng/µL in 5 µL nuclease-free water. A ligation reaction was set up in a sterile, nuclease-free 200 µL PCR tube as outlined in Table 3.2.

Table 3.2. Ligation of 3’ adapter reaction mixtures

Reagent	Volume (µL)
RNA 3’ Adapter (RA3)	1
3 µg Total RNA in nuclease-free water	5
Total Volume per Sample	6

The entire volume was mixed thoroughly by gently pipetting up and down 6 to 8 times followed by brief centrifugation. The reactions were then denatured at 70 °C for 2 minutes and immediately the tube was placed on ice. It is very important to place the tube on ice right after the 70 °C incubation to prevent secondary structure formation. The thermal cycler was then preheated to 28 °C and a new ligation mixture was prepared in another new sterile, nuclease-free 200 µL PCR tube as outlined in Table 3.3.

Table 3.3. New ligation reaction mixtures

Reagent	Volume (µL)
Ligation Buffer (HML)	2
RNase Inhibitor	1
T4 RNA Ligase 2, Deletion Mutant	1
Total Volume per Sample	4

The entire volume was mixed thoroughly again by gently pipetting up and down 6 to 8 times followed by brief centrifugation. Four µL of the mixture was then mixed with the entire volume from the earlier step to the same reaction tube. The total volume of the reaction should be 10 µL. The entire volume was mixed thoroughly followed by brief centrifugation. The tube was then incubated at 28 °C for 1 hour. After incubation, 1 µL of Stop Solution (STP) was added to the reaction tube and the mixture was gently pipette up and down for 6 to 8 times. The incubation was continued for another 15 minutes before immediately placed the tube on ice.

3.5.3.3 Ligation of 5’ adapter

The thermal cycler was preheated to 70 °C. The 5’ adapter was prepared by aliquoting 1.1 x N µL of the RNA 5’ Adapter (RA5) into a separate, nuclease-free 200 µL PCR tube, with N being the number of samples being processed for the experiment. The reaction tube was then incubated at 70 °C for 2 minutes and then immediately put on ice. The thermal cycler was then

preheated to 28 °C. The 5’ ligation mixture was then prepared as outlined in Table 3.4.

Table 3.4. Ligation of 5’ adapter reaction mixtures

Reagent	Volume (µL)
10 mM ATP	1.1
T4 RNA Ligase	1.1
Total Volume per Sample	2.2

An excess volume was prepared accounted for the loss during evaporation and to facilitate the pipetting of the proper volume at the next step. Two µL of the ligation mixture was then added to the 1 µL of the denatured RNA 5’ adapter and the reaction tube was thoroughly mixed by pipetting up and down for 6 to 8 times. The 5’ adapter mixture was then mixed with the entire volume from the “Ligate 3’ Adapter” tube by gently pipetting followed by brief centrifugation. The total volume of the reaction should be 14 µL. The reaction tube was then further incubated at 28 °C for 1 hour and immediately placed on ice after the incubation.

3.5.3.4 Reverse Transcription of Captured miRNAs

Reverse transcription followed by PCR was performed to create cDNA constructs based of the miRNA that were ligated with 3’ and 5’ adapters. This process selectively enriched those fragments with the adapter molecules on

both ends. The previous step resulted in a reaction volume of 14 µL. However, only 6 µL was used in the subsequent RT-PCR step, and the remaining of the 5’ and 3’ adapter-ligated RNA can be stored at -80 °C. Before start, the 25 mM dNTPs stock was diluted to 12.5 mM dNTPs using ultra-pure water in a separate, sterile nuclease-free, 200 µL PCR tube. The entire volume was mixed thoroughly by gently pipetting the mixture up and down, followed by brief centrifugation. The RT primer annealing mixture was prepared as outlined in Table 3.5.

Table 3.5. Reverse transcription primer annealing reaction mixtures

Reagent	**Volume (µL)**
5’ and 3’ Adapter-ligated RNA	6
RNA RT Primer (RTP)	1
Total Volume per Sample	7

The mixture was gently pipetted up and down for 6 to 8 times to ensure a thorough mixing, and then was briefly centrifuged. The reaction tube was then incubated at 70 °C for 2 minutes, before placed on ice. The thermal cycler was preheated to 50 °C. The RT-PCR reaction mixture was prepared as outlined in Table 3.6.

Table 3.6. Reverse transcription reaction mixtures

Reagent	Volume (µL)
5X First Strand Buffer	2
12.5 mM dNTP mix	0.5
100 mM DTT	1
RNase Inhibitor	1
SuperScript II Reverse Transcriptase	1
Total Volume per Sample	5.5

The mixture was mixed thoroughly by pipetting up and down, followed by a brief centrifugation. The RT-PCR reaction tube was then mixed with the 7 µL from the RT primer annealing mixture resulted in a total volume of 12.5 µL. The entire volume was mixed properly, centrifuged briefly and incubated at 50 °C for 1 hour before immediately placed on ice.

3.5.3.5 PCR Amplification

PCR Amplification reaction mixture was prepared in a new sterile, nuclease-free, 200 µL PCR tube placed on ice with each index added separately. The PCR Amplification reaction mixture was prepared as outline in Table 3.7.

Table 3.7. PCR amplification reaction mixtures

Reagent	Volume (µL)
Ultra Pure Water	8.5
PCR Mix (PML)	25
RNA PCR Primer (RP1)	2
RNA PCR Primer Index (RPI)	2
Total Volume per Sample	37.5

The reaction mixture was then mixed thoroughly by pipetting up and down 6 to 8 times, followed a brief centrifugation. The total volume of the reaction tube should be 50 µL. Subsequently, the PCR reaction was performed at 98 °C for 30 seconds, followed by 11 amplification cycles of denaturation at 98 °C for 10 seconds, annealing at 60 °C for 30 seconds and extension at 72 °C for 15 seconds. Final extension at 72 °C for 10 minutes and finally holds at 4 °C was also included in the PCR reaction. The reaction tube was kept at -20 °C before the next step.

3.5.3.6 Quality Check on the cDNA constructs

Each sample after the PCR amplification was subjected on a high sensitivity DNA chip (HS-DNA) (Agilent Technologies, USA) according to the manufacturer's instructions. Prior to checking, the cDNA constructs concentration were checked using Qubit dsDNA assay kit (Molecular Probes, USA). Briefly, 2 µL from the 50 µL PCR product was used for the Qubit dsDNA measurement. Two µL was directly mixed with 198 µL Qubit working solution, and the required low and high standards tubes were also prepared accordingly. The measurement of the standards and samples were then performed on Qubit 2.0 Fluorometer (Thermo Fisher Scientific, USA). The concentration of the DNA was calculated based on the Qubit value and multiply by the dilution factor. The concentration of cDNA constructs was adjusted to 4 ng/µL before loaded onto the HS-DNA chip. The dye concentrate and DNA gel matrix were first equilibrated to room temperature for 30 minutes. The dye concentrate was then vortex and spun down before 15 µL of it mixed with the gel. The complete gel-dye mix was placed on top of a spin filter and was centrifuged at 6000 rpm for 10 minutes. Load 9 µL of the gel-dye mix onto the DNA chip, followed by 5 µL of DNA marker, 1 µL of DNA ladder, and lastly 1 µL of the diluted DNA sample. The DNA chip which can accommodate up to 11 samples were then vortexed for 1 minute at 2400 rpm and before subjected to run on the Agilent 2100 Bioanalyzer. The sample wells were then detected by fluorescence, and electropherogram were created by the data analysis software for assessing DNA quality. An electropherogram with a flat baseline, all sample peaks were well resolved and appeared between the

lower and upper marker peaks indicated a successful HS-DNA run. The appearance of peaks at 145 to 160 bp indicated the successful ligation, reverse transcription and PCR amplification of the miRNAs and the samples can be preceded for subsequent purification step.

3.5.3.7 Purification of cDNA Constructs using Gel Electrophoresis

All the individual libraries constructed from the unique indices were pooled and gel purified together. The grouping for the gel pooling was determined based on the molarity (pmol/l) of the library ranged from 145 to 160 bp from the Agilent High Sensitivity DNA run result. Equal volumes of the library were combined and then the samples were loaded on the gel.

3.5.3.7.1 Preparation of 6% Polyacrylamide Gel (PAGE)

For the gel purification, Mini-PROTEAN Tetra Cell gel caster and electrophoresis system were used (Bio-Rad Laboratories, USA). The glass plates and spacers were first cleansed thoroughly with soap and water, rinsed with deionized water and lastly with ethanol to ensure there were no oils deposited on the working surfaces of the plates. The glass plates were set aside to dry. After dried, the glass plates with spacers were assembled in the gel caster. The gel solution with 6% polyacrylamide was prepared as outlined in Table 3.8.

Table 3.8. Preparation of 6% PAGE reaction mixtures

Reagent	Volume
30% Acrylamide/Bis (29:1)	2.4 mL
Deionized water	7.2 mL
5X TBE	2.4 mL
10% Ammonium persulphate (APS)	200 µL
TEMED	10 µL
Total Volume (approximate)	12 mL

The gel solution was prepared by mixing all the monomer solution as listed above except for the TEMED and 10% APS which were added last to initiate polymerization of the gel. The 1.0 mm comb was then inserted carefully into the gel as not to create any air bubbles trapped under the comb. The remaining gel solution was then layered onto the gel mould and any leakage of gel solution from the gel mould was observed. The acrylamide was allowed to polymerize for 45 to 60 minutes at room temperature. When the polymerization completed, the gel mould was wrapped with paper towel that had been soaked in 1X TBE, sealed with a plastic bag and stored in 4 °C up to a week. When ready for electrophoresis run, the gel was removed from the gel caster and inserted into the designated place in the mini tank. Running buffer, consisted of chilled 1X TBE buffer was added into the tank to the indicated level (550 mL for 2 gels) and the combs were carefully pulled out from the gels. The wells were then flushed twice with 1X TBE with a Pasteur pipette.

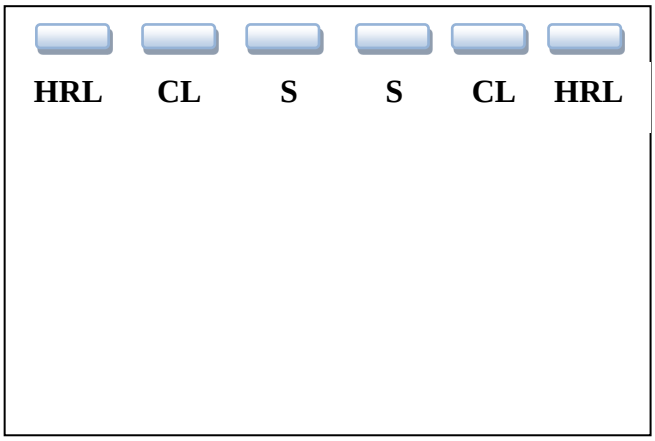
3.5.3.7.2 Run Polyacrylamide Gel Electrophoresis

Two µL of Custom ladder was mixed with 2 µL of DNA loading dye while 1 µL of high resolution ladder was mixed with 1 µL of DNA loading dye. All the amplified cDNA constructs that had been grouped based on their molarity were mixed with 10 µL of DNA loading dye. The total volume of the mixture should not exceed 50 µL. Two µL of mixed custom ladder-loading dye was loaded in two separate wells on the 6% PAGE gel. Two µL of mixed high resolution ladder-loading dye was also loaded in a different well. 50 µL of pooled amplified cDNA constructs mixed with loading dye were loaded on the into two wells 6 % gel with 25 µL each. The ladder and sample arrangement on the gel is illustrated in Figure 3.4. The electrodes were connected to a power pack, power was turned on, and the electrophoresis was run for 50 minutes at 150 V.

3.5.3.7.3 Gel Extraction of Purified cDNA Constructs

The gel was removed from the apparatus and stained with 3X Gelred nucleic acid (Biotium, USA) for 15 minutes before visualised with UV Bioimaging system (UVP, USA). Using a clean razor blade for each sample, the sharp band located between the custom ladder 145 and 160 bp, was excised. Both lanes containing the same samples were combined into one slice and placed into the gel breaker tube attached to a sterile 2 mL microcentrifuge tube. The tube was then centrifuged for 2 minutes at room temperature to move the gel through the holes made in the gel breaker tube into the 2 mL tube.

Ensure that the gel has moved through the holes into the bottom tube. Discard the gel breaker tube. 300 µL of ultrapure water was then added to the gel debris in the 2 mL tube and the tube was left rotating in the shaker incubator at room temperature for 2 hours. The eluate and the gel debris were then transferred to the top of a 5 µm filter tube. The filter tube was centrifuged for 10 seconds at 600 g and 2 µL of glycogen, 30 µL of 3M NaOAc, and 975 µL of pre-chilled 100% Ethanol were then added to the eluate. The mixture was then immediately centrifuged at 20, 000 g for 20 minutes on a benchtop microcentrifuge set at 4 °C. The supernatant was discarded, leaving the pellet intact. The pellet was carefully washed with 500 µL of 70% Ethanol at room temperature, before centrifuged again at 20,000 g for 2 minutes. The supernatant was discarded, and the intact pellet was left dried with open lid in a 37 °C heat block for 10 minutes. After that, the pellet was resuspended in 10 µL 10 mM Tris-HCl, pH 8.5 and kept in -20 °C until further use.



Legends: -

HRL: High Resolution Ladder

CL: Custom Ladder

S: Pooled cDNA Sample

Figure 3.4. Purification of cDNA constructs containing the miRNAs arranged on a 6% PAGE gel. The markers were loaded on both edges of the gel.

3.5.3.8 Library Validation and Mixing

The quality and concentration of each final library was then assessed using the Agilent Bioanalyzer DNA high sensitivity chip (Agilent Technologies, USA) according to the protocol mentioned in Section 3.5.3.6.An electropherogram with a flat baseline, and a peak that appeared between 120 bp to 170 bp within the lower and upper marker peaks indicated a successful HS-DNA run. The peak with average size of 140 to 145 bp indicated a successful DNA library purification from the PAGE gel. The pooled of DNA library samples were then diluted to 2 nM in 5 µL each with nuclease-free water. Then, the 2 nM libraries were combined at equimolar ratios into a single multiplexed library. The remaining DNA samples were stored in Tris-HCl with 0.1% Tween-20 in -20 °C for long term storage. Finally, 2 µL from the single multiplexed library was subjected to the DNA measurement using Qubit dsDNA HS assay kit (Molecular Probes, USA)**.** DNA concentration > 0.2 ng/µL indicated a DNA library yield suitable to be used for sequencing. The HiSeq 2500 Illumina Sequencer for TruSeq Small RNA Sample Preparation available in High Research Impact Building in University of Malaya with single indexing and rapid run was used for the cluster generation.

3.6 Bioinformatics Analysis of the miRNA Library Sequencing Data

CLC Genomic Workbench Version 7.0 (Qiagen Bioinformatics, Hilden, Germany) installed in the workstation located in the Bioinformatics Laboratory

of Institute of Bioscience, Universiti Putra Malaysia was used as the analysis pipeline. The overview of the analysis workflow is outlined in Figure 3.5.

3.6.1 Deposition of miRNA NGS Raw Data in Public Databases

miRNA datasets generated in this study were deposited at Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/info/seq.html). Briefly, the information required in the metadata spreadsheet was filled up and together with the original unedited files from the Illumina pipeline in fastq format were uploaded onto the FTP deposit area via free client software, FileZilla. The FTP login information was connected to the host ftp-private.ncbi.nlm.nih.gov under the "fasp" directory and with "geo" as the username and "33%9uyj_fCh?M16H" as the password.

3.6.2 FastQC for High Throughput Data

A quality control tool was used to check on the quality of the raw sequence data before doing any further analysis. This tool reads raw sequence data from high throughput sequencers and produces a report of the overall quality of the run, any potential problems or biases. The FastQC tool was downloaded at (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and run as an interactive graphical application in a Java runtime environment (JRE) build 1.8.0_60. To launch FastQC, the run_fastqc bat file in its installed

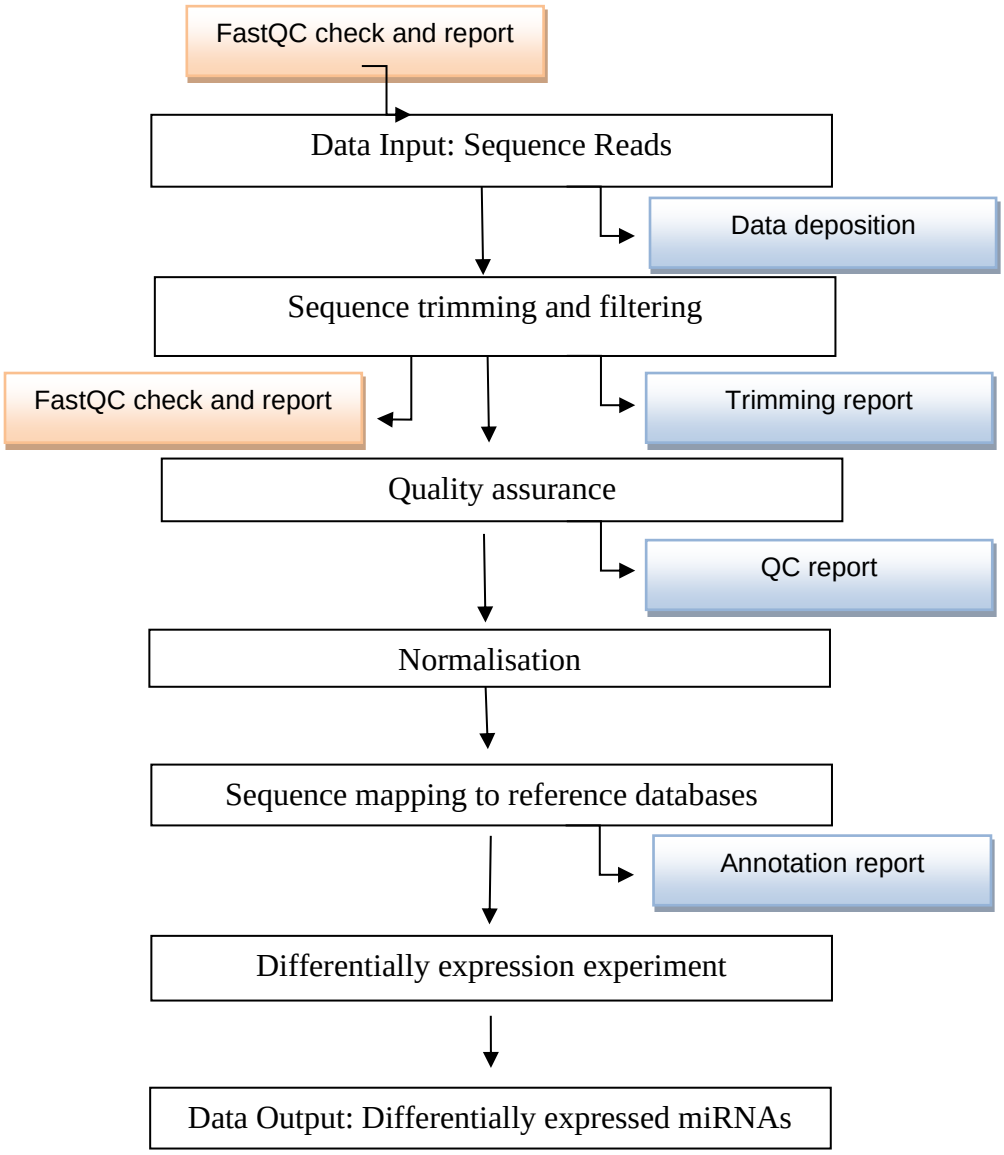
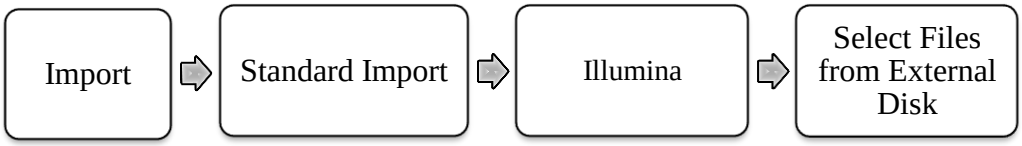


Figure 3.5. Schematic overview of the bioinformatics workflow used to analyse the miRNA transcriptomic data for the identification of differentially expressed miRNAs.

location was double clicked. Unzipped one of the raw sequencing data folder and the fastq file was imported into the FastQC. The fastq file started to be processed once it was uploaded onto the interface. The quick analysis provided an overview of the fastq file which includes basic statistics, per base sequence quality, per base sequence quality scores, per base sequence content, GC content, sequence length distribution and also overrepresented sequences. FastQC was also performed on the trimmed data (removal of adapters) and a good quality data should have Q30 base sequence quality scores across all bases and no overrepresented sequences found.

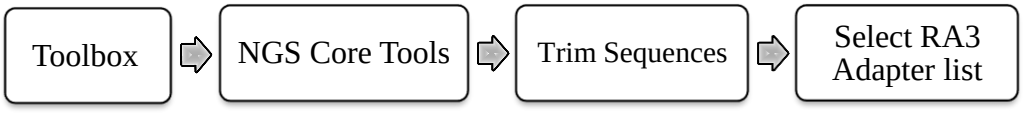
3.6.3 Standard Import of Raw Sequencing Data

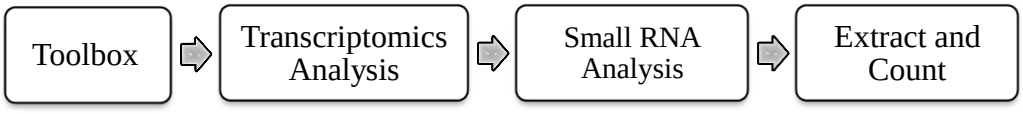
Before analysis can be performed, all the data has to be imported into the workbench using standard Illumina data importer. The fastq file for import (files name end with fast.gz) was selected. The 'Discard quality scores and discard read names' checkboxes were checked as information about quality scores and read names were not used in analysis. This in turn helped to save the disk space when importing the data. 'NCBI/Sanger or Illumina Pipeline 1.8 or later' was chosen as the preferred quality scores pipeline. The data import was initiated by choosing the tools in CLC Genomics as follows:



3.6.4 Trimming Adapters and Counting the Reads

In the next step, the imported data was trimmed to remove the partial adapter sequences and subsequently the number of copies of each of the resulting miRNAs was counted. Before that, the Illumina Truseq Small RNA 3' Adapter (RA3) trimming list was downloaded from http://www.clcbio.com/files/tutorials/small_RNA_analysis/Illumina_3_RNA_Adapter_Trim_List.csv. The sample fastq file to be trimmed and the adapter list were selected and subjected to Transcriptomics Analysis – Small RNA Analysis- Extract and Count. The minimum length of the tags was set to 17 and the maximum length of the tags was set to less and equal to 25. The other options appeared in the 'Extract and Count' dialog box was set to the default values. The checkboxes of 'Create sample' and 'Create report' were checked in the output option and a report containing the trim summary, read length before/after trimming, for each of the input file was created. The trimming and reads counting were performed using tools in the CLC Genomics as follows:



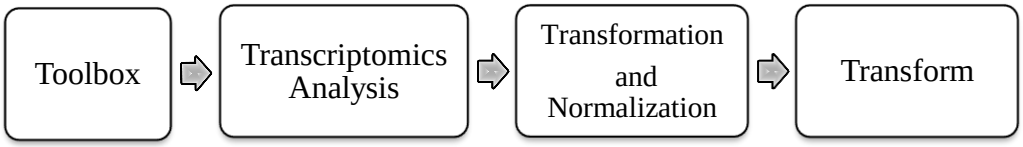


3.6.5 Quality Control

A number of quality control analyses were performed to inspect the overall distribution, variability and similarity of the sets of expression values in samples. Samples of poor quality and outlying samples were excluded.

3.6.5.1 Data Transformation and Normalization

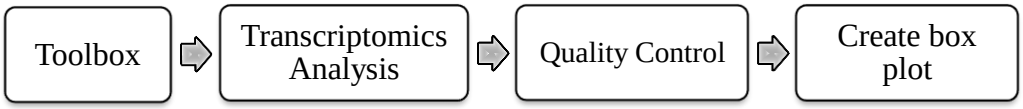
The original expression values of all the samples were transformed and normalized in order to ensure all the samples were comparable for further data analysis. To transform the sets of original expression values, the tools in CLC Genomics were chosen as follows:



Log2 transformation was chosen as the logarithm. Quantile normalization was chosen as a way of normalization the sets of expression values by calculating the empirical distribution of the sets of expression values for the samples. The new transformed and normalized values were then added to the new columns of the analysed expression values, the original expression values were not changed. The outcome of the transformation and normalization was then saved into the final output file.

3.6.5.2 Distribution Analysis Using Box Plots

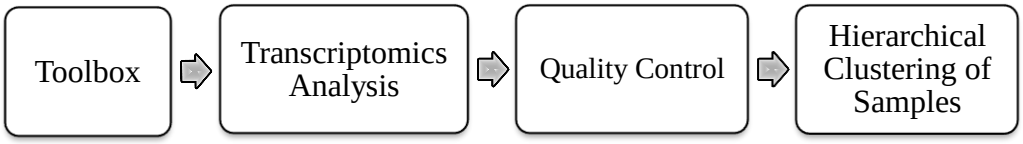
The overall distribution of the sets of expression values in samples of the study was visualised using a box plot. The box plot was created as follows:



The number of samples were then selected and included for the generation of the box plots. The distributions of the samples' values were presented by a line at the centre, the middle part represented by a box, and the tails of distribution represented by the tails. Differences found in the overall distributions of the samples were fixed with sample normalization as described in Section 3.6.5.1. Normalized data sets should have distributions with the equally sized boxes and whiskers.

3.6.5.3 Hierarchical Clustering of Samples

A hierarchical clustering of samples is a tree representation of their relative similarity. The clustering was created as follows:

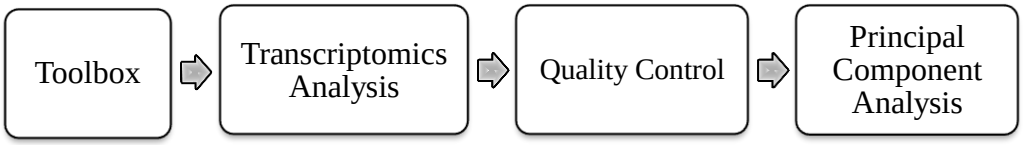


The number of samples were then selected and included for the generation of the clusters. The 1-Pearson correlation and single linkage were chosen as the

distance measures and the cluster linkage respectively. The outcome of the clustering was then saved to the experiment and can be viewed by clicking "Show Heat Map" on the final output file.

3.6.5.4 Principle Component Analysis

A principle component analysis (PCA) is a mathematical analysis that used to visualize the variation in the datasets. The PCA analysis was started as follows:

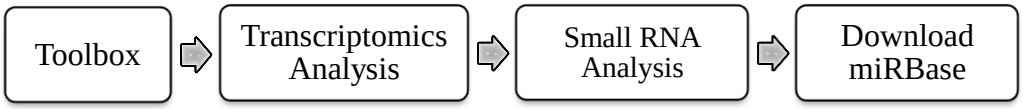


The number of samples were then selected and included for the generation of the PCA plot. The projection of the samples onto the two-dimensional space spanned by the first and second principal component of the covariance matrix was created. The group relationships were indicated by different colour. Outlying samples or samples that had been wrongly assigned to a group were identified.

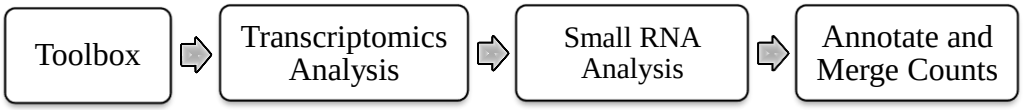
3.6.6 Annotating and Merging Small RNA Samples

After all the quality assessments on the datasets were completed, annotation of the small RNA was performed. Prior to that, two sources for the annotation; first, miRBase to identify known miRNAs and second a set of other

known non-coding RNAs were downloaded onto the CLC Genomics Workbench. The first annotation file on the information of the mature regions on the precursor miRNAs was downloaded from the miRBase-Release 21 via integrated tools as follows:



The second annotation file consists of a list of *Homo sapiens* non-coding RNAs fromENSEMBL (The Ensembl genome database project) was downloaded as a fasta filefrom ftp://ftp.ensembl.org/pub/release-75/fasta/homo_sapiens/ncrna/Homo_sapiens.GRCh37.75.ncrna.fa.gz. Besides the known small RNAs were annotated in the sample, the variants of the same small RNA were also merged to get a cumulative count. The trimmed reads which were generated earlier were then subjected to annotation and sample merging using the integrated tools as follows:

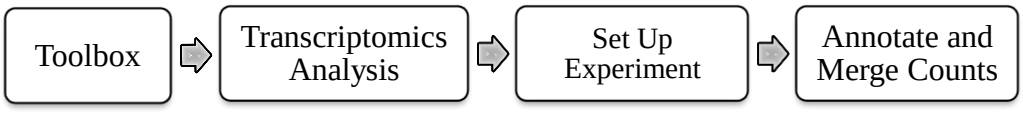


An open dialog box was opened where the small RNA samples to be annotated were selected. Several samples of the same group were included for this process. The annotation resources to be used were defined by selecting the miRBase-Release 21 that was downloaded earlier. Another source to be used was defined by selecting the Homo_sapiens. GRCh37.75. ncrna file that was imported. MiRBase was set as the highest priority for annotation purpose.

Since the miRBase file contains a list of precursor sequences from different species, *Homo sapiens* was selected in the next dialog box as the annotation source. *Mus musculus* species was also included to be the second in the list, since there may be some miRNAs that are new to human but have an ortholog in mouse. The other settings in the dialog box were set as default. The other options were left checked except the unannotated samples/experiments in the next dialog box. One of the important parameters that were included was the 'created grouped sample, grouping by Mature' in the output of the analysis, where the two precursor variants of the same mature 5'miRNA were merged as one entity. The results were then saved in the output file.

3.6.7 Experimental Design

In order to analyse the differential expression, the experiments involving the analysed samples were set up on the CLC Genomics Workbench. An experiment consisted of the sets of samples which were grouped and the relationships between the samples were defined. A statistical analysis was then performed to investigate the differential expression between the groups. An experiment was set up as follows:

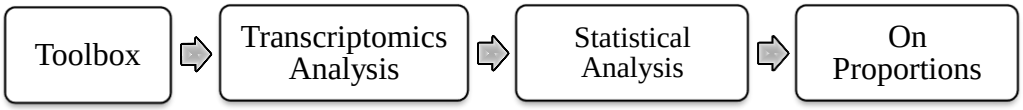


Unpaired two-group experiment was defined as the number of groups used for comparison. The group names were then assigned accordingly in the

designated box, e.g. Group1: Parental and Group 2: Spheroid. Then, the individual samples belong to the experiments were then defined.

3.6.8 Statistical Analysis and Volcano Plots

The Kal's Z-test, a proportions-based statistical test was used for the differential expression analysis as follows:



Four columns were added to the experimental table for each pair of groups that were analysed when subjected to Kal's Z-test. The first column 'Proportions difference' consisted of the difference between the proportion in group 2 and the proportion in group 1. The second column 'Fold Change' showed how many times bigger the proportion in group 2 relative to that of group 1. Positive value of the fold change was obtained if the proportion in group 2 is bigger than that in group 1 and vice versa. The third column 'Test Statistic' displayed the value of the test statistic and the last column 'p-values' showed the two-sided p-value for the test. The p-values were then corrected by choosing Bonferroni corrected and FDR corrected before the results were saved. To further inspect the results of the statistical analysis, the volcano plots were then examined by clicking the 'Show Volcano Plot' button at the bottom of the experiment table view. Since the plots used the p-values and the mean

difference produced by the statistical analysis, the plots were only available once the statistical tests had been performed on the experiment.

3.6.9 Sorting and Filtering the Experimental Tables

After the experiments were set up, completed with statistical analysis, the differential expression lists were ready to be sorted and filtered according to certain features. The columns under the 'Experiment' header were used for the filtering purpose. Volcano plots were examined to confirm the filtering analysis. The filtering features were performed as follows:

Significant	Significant, UP -regulated	Significant, DOWN-regulated
• Kal's Z-test Proportions Fold Change > 2 • Kal's Z-test FDR-corrected *p*-value < 0.05	• Kal's Z-test Proportions Fold Change > 2 • Kal's Z-test FDR-corrected *p*-value absolute value < 0.05	• Kal's Z-test Proportions Fold Change < -2 • Kal's Z-test FDR-corrected *p*-value absolute value < 0.05

After the lists of miRNAs were filtered according to the filtering features, a subset of the experiment was created. The relevant features (rows) were first highlighted and the subset of the experiment was created by clicking the 'Create Experiment from Selection' button at the bottom of the experiment table. This new list with the desirable features was then exported out as a Microsoft Excel document to be used for analysis in the next section.

3.7 Bioinformatics Analysis of the Differentially Expressed miRNA Lists

Before further analyses were performed on the differentially expressed miRNA lists, a Venn diagram (http://www.bioinformatics.lu/venn.php) was constructed to analyse the differentially or the commonly expressed miRNAs between the samples. After that, further bioinformatics analyses were carried out using web-accessible programs such as Cytoscape Version 3.3 (http://www.cytoscape.org/) and the online Database for Annotation, Visualization and Integrated Discovery (DAVID) program (https://david.ncifcrf.gov/). Cytoscape was used to predict the miRNA-target genes and also to visualize the complex underlying networks between the miRNAs and their predicted target genes. On the other hand, DAVID program was used for functional annotations (Gene Ontology, GO) and to elucidate the enriched pathways based on Kyoto Encyclopaedia of Genes and Genomes (KEGG). The overview of the analysis workflow is as follows in Figure 3.6.

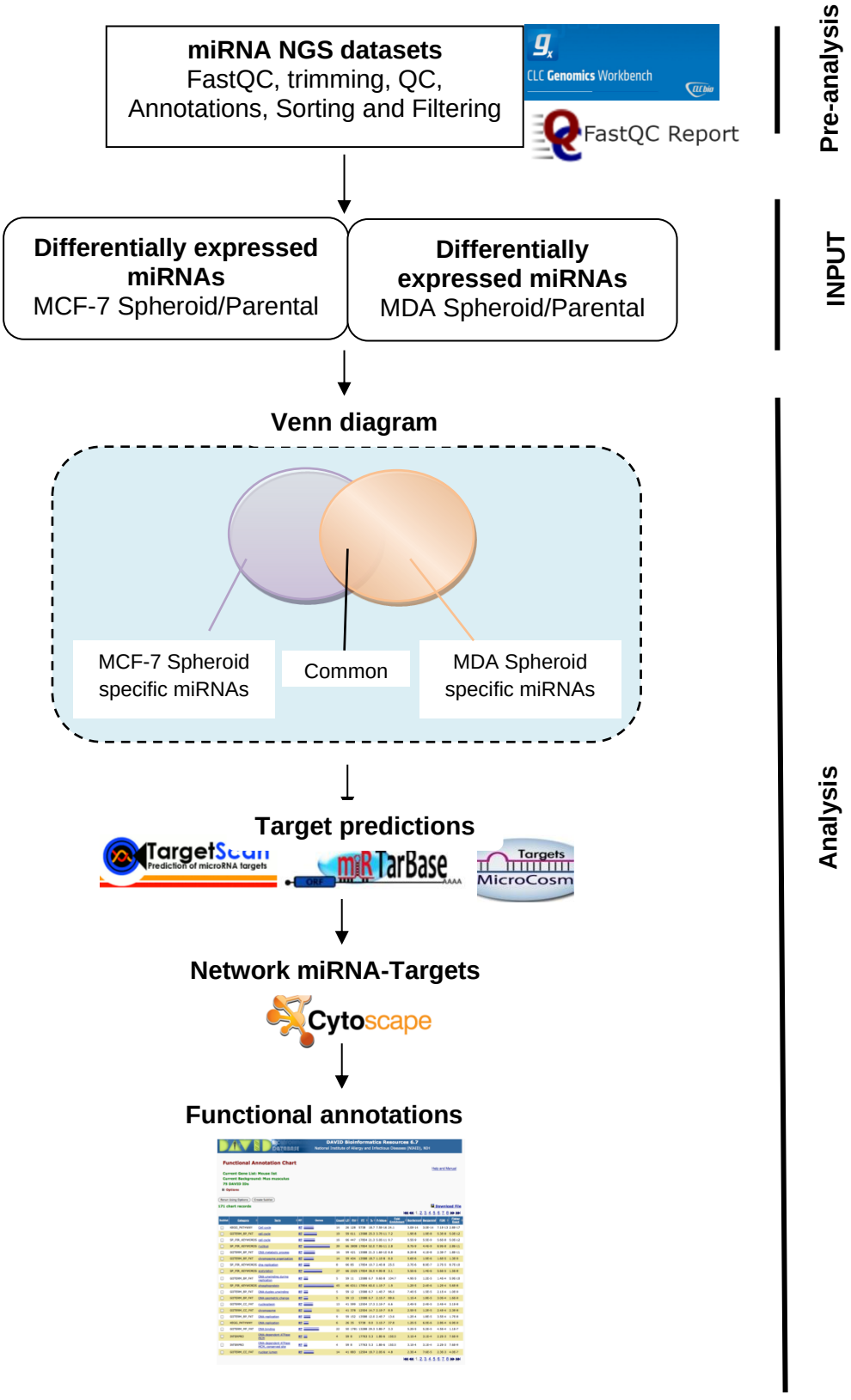


Figure 3.6. Outline of miRNA transcriptomic integrated analysis using different bioinformatics tools

3.7.1 miRNA Target Predictions and Network Visualization Using CyTargetLinker

Cytoscape Version 3.3, which is an open source software platform was first downloaded from http://www.cytoscape.org/ and installed into the computer directory. CyTargetLinker, which is one of the open source applications developed for the visualization of the biological networks with regulatory interactions networks to be used in Cytoscape 3.3 was downloaded from the Cytoscape App Store (http://apps.cytoscape.org/apps/cytargetlinker). The regulatory interactions networks (RegINs) used in CyTargetLinker were derived from an online database, stored in XGMML (the eXtensible Graph Markup and Modeling Language) format, which is supported by Cytoscape. Each RegINs consists of two nodes, a source (regulatory component) and target biomolecule, connected through one directed edge. The collection of RegINs for Homo sapiens under the microRNA-target interactions were downloaded from the CyTargetLinker website (http://projects.bigcat.unimaas.nl/cytargetlinker/regins/).

A biological network containing the top ten deregulated differentially expressed miRNAs was created in Cytoscape Version 3.3. The Microsoft Excel data file was first generated in a table format containing two columns of information, i. the microRNA name and ii. the corresponding miRBase accession number. The data file was then imported into Cytoscape interface as follows:

In the ‘Import Network and Edge Attributes from Table’ dialog box, the column containing the microRNA names were selected as the ‘source interaction’ and the header of the column were transferred as ‘attribute names’. The miRBase accession number present in the same data file was then defined as ‘Source Node Attribute’. The resulting biological networks of the microRNAs were then appeared in the Cytoscape user interface. The CyTargetLiner app which was downloaded earlier was then activated as a plugin in the Cytoscape as follows:

The CyTargetLinker selection panel appeared after the plugin was loaded onto the Cytoscape. Human microRNAs network was selected as the user network followed by miRBase accession numbers as the network attribute. The directory containing the Homo sapiens specific RegINs was then selected, and they were miRtarbase-3.5-hsa.xgmml, targetscan-hsa-2012-12-05.xgmml and microcosm-hsa-202-12-05.xgmml. The name of the resulted extended biological network was then renamed. The direction of the interaction was chosen to be ‘both’ (interactions from source to target (adding target nodes) and interactions from target to source (adding regulatory node). The extended

network was then visualised in the user interface whereby the network microRNAs and target genes were defined as red circles and pink pentagons, respectively. The target genes list predicted using the three prediction programs was the saved as an output list and subjected to further functional annotations using web-based tool DAVID.

3.7.2 Gene Set Enrichment and Pathway Analysis Using DAVID

Gene set enrichment analysis was performed to further elucidate the functions of the predicted target genes by using web-based tool DAVID. The analysis was started by accessing the DAVID website https://david.ncifcrf.gov/. In the DAVID interface, the predicted target gene list was submitted onto the designated box 'Upload Gene List' on the left panel. The DAVID Gene Functional Classification Tool was then selected for the submitted genes list. The underlying themes of the significantly differentially expressed predicted genes for the biological processes (BP), molecular functions (MF), cellular components (CC) and enriched pathways using Kyoto Encyclopaedia of Genes and Genomes (KEGG) were assessed. The modified Fisher Exact p-values were then calculated using the software algorithm, with smaller *p*-values to be considered strongly enriched in the annotation categories. The resulting p-values were then sorted, and the output results were then saved.

3.8 miRNAs Validation by real-time PCR

A total of twelve differentially expressed miRNAs that were > 2 fold change represented specific pathways found in MCF-7 and MDA-MB-231 spheroid-enriched pathways were identified for further validation. The accession number mature sequences of the selected miRNAs were retrieved from the miRBase database (http://www.mirbase.org/). All the primers were synthesised using the miRNA LNA PCR primer sets (Exiqon, Vedbæk, Denmark) and are listed in Table 3.9. Both PCR amplification primers (forward and reverse) were miRNA specific and optimized with locked nucleic acid (LNA) technology enabled high sensitivity and specificity during the amplification of short RNA targets.

The miRNA validation by real-time PCR utilized miRCURY LNA Universal RT microRNA PCR System (Exiqon, Denmark) consisted of two-part protocol involving, i. first-strand cDNA synthesis and ii. Real-time PCR amplification. The incubation and PCR amplification were performed in a CFX96 Touch Real-time PCR Detection System (Bio-Rad Laboratories, USA). In brief, total RNAwith retention of miRNAs for each sample was first adjusted to a concentration of 50 ng/µL using nuclease-free water. The reaction buffer was then thawed on ice for 15 minutes, flicked, and spun down. The first-strand cDNA synthesis reaction mixture was prepared as outlined in Table 3.10.

Table 3.9: **List of miRNA primers used in real-time PCR**

Name	Sequence accession number	Target Sequence
hsa-miR-19a-3p (Reference)	MIMAT0000073	UGUGCAAAUCUAUGCAAAACUGA
hsa-miR-200a-3p (Reference)	MIMAT0000682	UAACACUGUCUGGUAACGAUGU
hsa-miR-15b-5p	MIMAT0000417	UAGCAGCACAUCAUGGUUUACA
hsa-miR-34a-5p	MIMAT0000255	UGGCAGUGUCUUAGCUGGUUGU
hsa-miR-148a-5p	MIMAT0004549	AAAGUUCUGAGACACUCCGACU
hsa-miR-628-5p	MIMAT0004809	AUGCUGACAUAUUUACUAGAGG
hsa-miR-196b-5p	MIMAT0001080	UAGGUAGUUUCCUGUUGUUGGG
hsa-miR-126-5p	MIMAT0000444	CAUUAUUACUUUUGGUACGCG
hsa-miR-760	MIMAT0004957	CGGCUCUGGGUCUGUGGGGA
hsa-miR-30c-5p	MIMAT0000244	UGUAAACAUCCUACACUCUCAGC
hsa-miR-136-5p	MIMAT0000448	ACUCCAUUUGUUUUGAUGAUGGA
hsa-miR-204-5p	MIMAT0000265	UUCCCUUUGUCAUCCUAUGCCU
hsa-miR-181a-5p	MIMAT0000256	AACAUUCAACGCUGUCGGUGAGU
hsa-miR-205-5p	MIMAT0000266	UCCUUCAUUCCACCGGAGUCUG

Table 3.10. cDNA synthesis reaction mixtures

Reagent	Volume (µL)	Final concentration
Reaction buffer (5X)	2.0	1X
Nuclease-free water	5.0	-
Enzyme mix (10X)	1.0	1X
Template total RNA (50 ng/ µL)	2.0	100 ng
Total Volume per Sample	10.0	

In brief, a polyA tail was added to the mature miRNA template by poly (A) polymerase and reverse transcribed to cDNA using a polyT primer with 3' degenerate anchor and a 5' universal tag. The cDNA template was then amplified using microRNA-specific and LNA-enhanced forward and reverse primers. The two processes were outlined in the workflow show in Figure 3.7. To synthesise first-strand cDNA, the 10 µL reaction tube was incubated at 42 °C for 60 minutes before heat-inactivated for 5 minutes at 95 °C in the thermal cycler. Synthesised cDNA was then subjected to real-time PCR amplification using SYBR green fluorescent dye detection system. Each reaction was performed in a final volume of 10 µL containing 1:1 ratio of diluted cDNA and the PCR master mix as outlined in Table 3.11.

Table 3.11. Real-time amplification reaction mixtures

PCR components	Volume (µL)	Final concentration
PCR Master mix (2X)	5.0	1X
PCR Primer mix	1.0	
Diluted cDNA template (1:40)	4.0	40X diluted
Total Volume	10.0	

Negative control was included for the PCR amplification by substituting the cDNA template with nuclease-free water. The amplification profile was denaturation at 95 °C for 10 minutes, followed by 40 amplification cycles at 95 °C for 10 seconds, and 60 °C for 1 minute. The ramp-rate cooling from 95 °C to 60 °C was set at 1.6 °C/seconds. The specificity and identity of the amplification reaction was analysed using the melt-curve analysis by slowly heating the PCR products from 70 °C to 95 °C in 0.5 °C increment. All reactions were run in triplicates and the procedure confirmed to the MIQE checklist. After the reaction, the threshold cycle (Ct) values were determined using the comparative Ct method for relative quantification. The Ct value is the value obtained from the intersection between an amplification curve and a threshold line generated from the number of PCR cycles required for the fluorescent signal to exceed the background level. The data were normalised to the reference miRNAs based on geNorm algorithms.

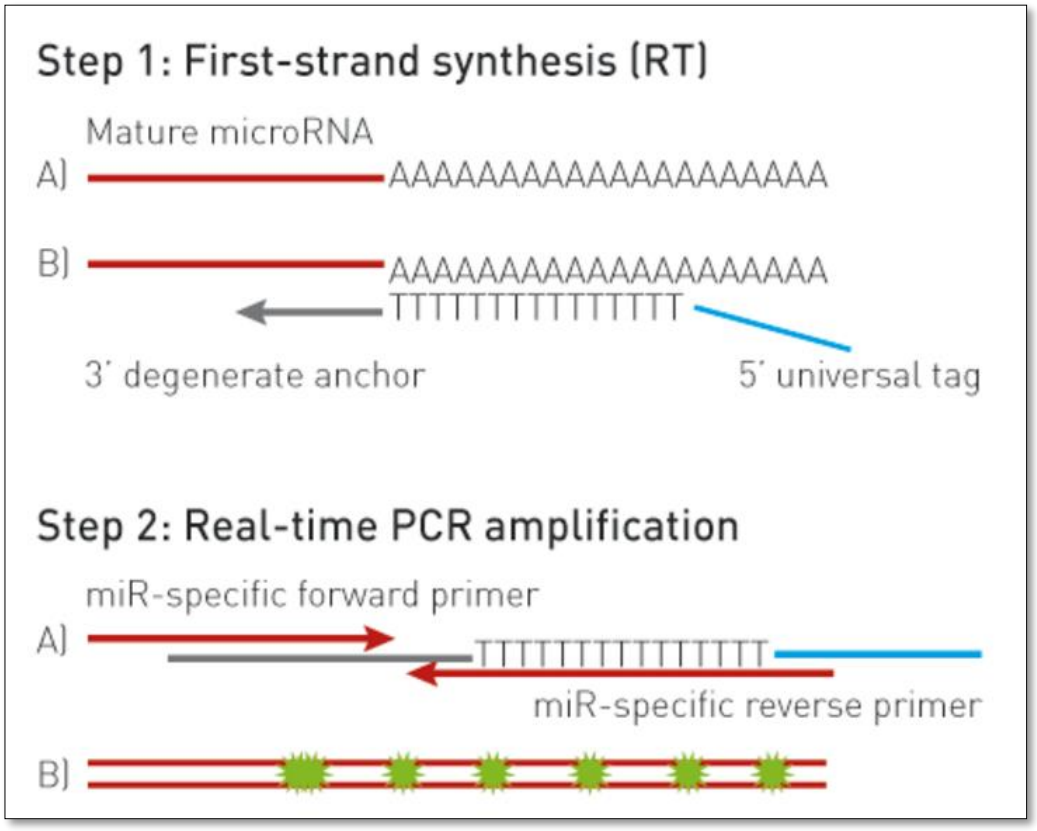


Figure 3.7. Schematic outline of the miRCURY LNA Universal RT microRNA PCR System (Alvarez M.L., 2014).

3.9 Statistical Analysis

Independent sample one-way analysis of variance (ANOVA) and student independent t-test via SPSS Version 17 (Statistical Package for Social Sciences) was used to determine the statistical significance of the data between the spheroid cells and parental cells. The results were considered to be statistically significant at a probability level of *p*-value <0.05. Statistical analysis of miRNA profiling NGS data was performed in the CLC Genomics Workbench Version 7 using Kal's Z test which is a proportions-based test used for differential expression analysis between the NGS datasets. Two-fold expressions were considered as the cut-off level. Functional annotations for the Gene Ontology and KEGG pathways were statistically analysed using the modified Fisher Exact p-values in DAVID program. The smaller the *p*-values, the data were considered to be stronger enrichment in the annotation categories. Real-time PCR results were reported as mean of log 2-fold change ± standard deviation (SD). Descriptive analysis was used to provide qualitative features about the observations that had been made in this study.

CHAPTER 4

RESULTS AND DISCUSSION PART 1: GENERATION AND CHARACTERIZATION OF SPHEROID CELLS ENRICHED WITH CANCER STEM CELLS PROPERTIES

Part of this chapter has been published in:

Boo, L., Ho, W. Y., Ali, N. M., Yeap, S. K., Ky, H., Chan, K. G., Yin, W. F., Dilan, A. S., Liew, W. C., Tan, S. W., Ong, H. K., Cheong, S. K. (2016). MiRNA transcriptome profiling of spheroid-enriched cells with cancer stem cell properties in human breast MCF-7 cell line. International Journal of Biological Sciences, 12(4), 427–445.

and

Boo, L., Ho, W. Y., Ali, N. M., Yeap, S. K., Ky, H., Chan, K. G., Yin, W. F., Dilan, A. S., Liew, W. C., Tan, S. W., Cheong, S. K., Ong, H. K. (2017). Phenotypic and microRNA transcriptomic profiling of the MDA-MB-231 spheroid-enriched CSCs with comparison of MCF-7 microRNA profiling dataset. PeerJ, *5*, e3551.

4.0 Results and Discussion Part 1: Generation and characterisation of spheroid cells enriched with cancer stem cells properties

4.1 Introduction

The initial part of the study focused on the establishment and characterisation of the spheroid-enriched cancer stem cells (CSCs) models from both cell types, and also highlighting other features, which have not been previously reported for spheroid cells. The CSCs models of tumour progression suggest that tumours contain a subset of cells that have self-renewal properties and give rise to differentiated progeny (Najafi, Farhood et al., 2019). These subsets of cells are hypothesised to reside in their microenvironment niche within the tumour and are the only cells that can maintain the tumour indefinitely due to their self-renewal ability (Hu, Mirshahidi et al., 2019). Their differentiation potentials were measured by the ability to give rise to the phenotypically diverse cancer cell population within the tumour. This population of cells are also responsible for the tumour initiation, maintenance and progression (Ayob and Ramasamy, 2018). As such, being able to isolate and characterise this subpopulation of cells from tumours remain a challenge.

Spheroid enrichment technique, based on the ability of CSCs to survive in a serum-free environment, was used for the enrichment of CSCs *in vitro*. In this study, MCF-7 and MDA-MB-231, the most established breast cancer cell lines, each representing an unique molecular subtype, were used to model

breast cancer. Recent analysis using gene expression microarray platform has divided breast cancer into four major molecular subtypes; luminal A, luminal B, triple negative/basal-like and HER2 type, suggesting that there might exist different origin of the different subtypes of tumours (Kumar, Prasad et al., 2018). As such, the high primary tumour heterogeneity in breast cancer poses another limitation for using this method for CSCs isolation. Besides, in most laboratories, primary breast tumour cells derived from patient's tissues often lose their native *in vivo* characteristics after being cultured in a tissue culture environment during the first few passages. Additionally, these cells have a limited lifespan, restricted furthered downstream analysis involving the same cells. Therefore, cell lines, which are now widely available through commercial cell banks, are the more preferred choice as they are homogenous in culture, overcoming the heterogeneity issues observed in clinical tumours (Begley and Ellis, 2012). Another major benefit of using cell lines is the indefinite lifespan on the cultured cells, allowing a continuous supply of cell population in cancer research. Questions have also been raised on how representative the immortalised breast cancer cell lines to their breast cancer tissue of origins. Continuous cell lines which lack cell-cell interaction and other functions based on the tumour environment context have also often been associated with genotypic and phenotypic drifting (Hynds, Vladimirou et al., 2018). Despite these limitations, breast cancer cell lines still prove to be valuable experimental tools and information derived from these has benefited the translational breast cancer research. Extensive research has also shown that cell lines continue to share many genetic and molecular features from which they were originated (Lacroix and Leclercq, 2004). One of the examples was

the use of CSCs derived from MCF-7 cells for the determination of the efficacy of sulforaphane, a compound from broccoli sprouts, which pave the way for the subsequent clinical investigations for breast cancer chemoprevention (Li, Zhang et al., 2010). In our research, MCF-7 and MDA-MB-231 breast cancer cell lines, representing two breast cancer molecular subtypes were used as the starting material. To our knowledge, breast cancer research is often modelled using these two established breast cancer cell lines particularly in the evaluation of stem-like cancer cells subpopulations (Burdall, Hanby et al., 2003, Holliday and Speirs, 2011). Since their first development, research has accelerated rapidly on both of these cell lines with nearly 32,176 scientific publications that used MCF-7 cell line and about 16,905 for MDA-MB-231 cell line. In our study, the cultured cell lines were negative for mycoplasma contamination, which is imperative to ensure all the obtained data were reliable and reproducible. Cells contaminated with mycoplasma have been known to have altered growth rates, morphology and metabolism, which could be regarded as a different cell line entity, a detrimental effect to research (Gedye, Cardwell et al., 2016). Therefore, mycoplasma-free cultures are pre-requisite for the generation of reliable data in the biomedical research field (Kim, Shin et al., 2019).

4.2 Experimental Design

To generate and characterise the spheroid cells from both cell lines, the experiments were performed as shown in Figure 4.1.

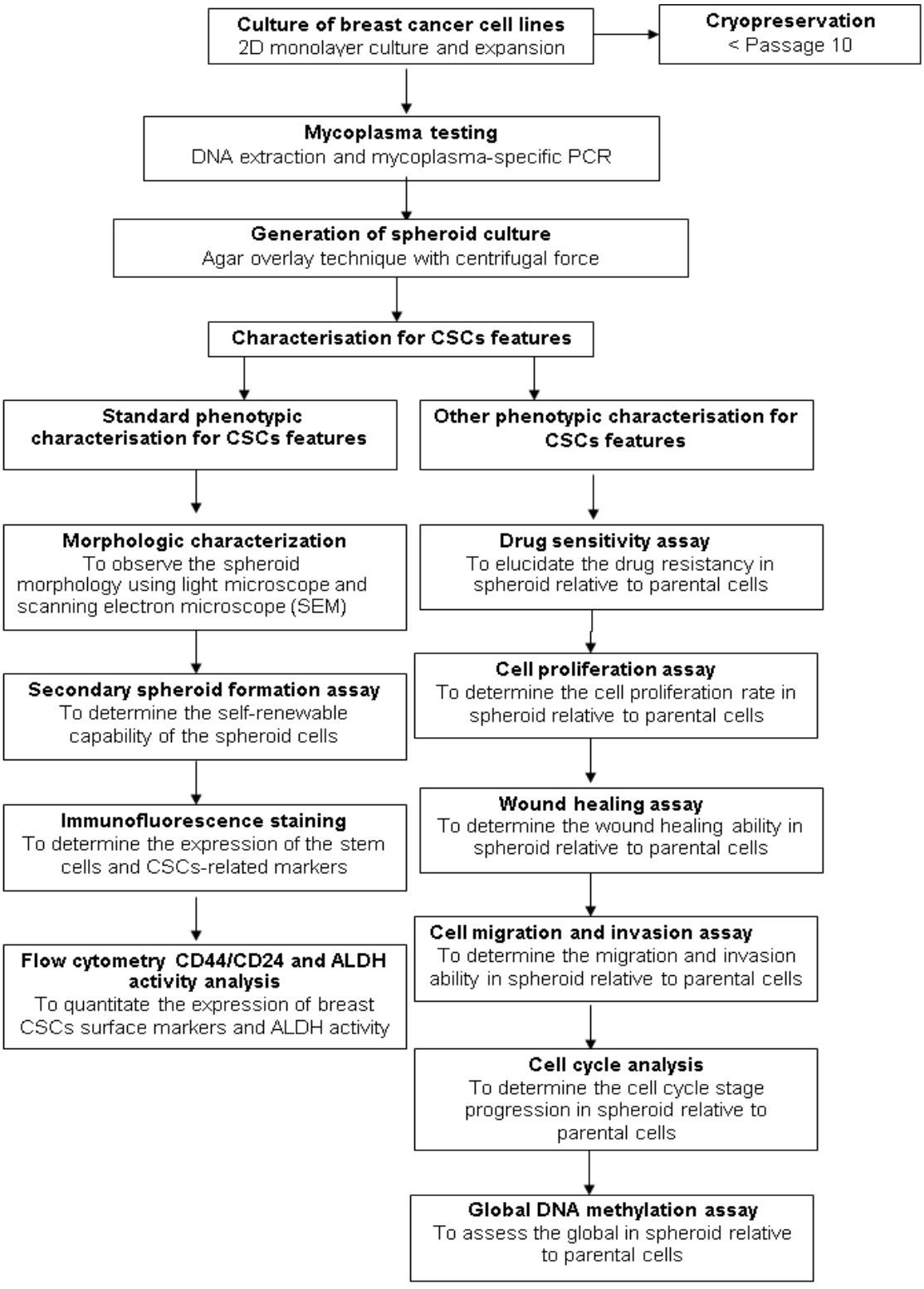


Figure 4.1. Experimental design of Part 1. The flow chart outlines the overall methodology for the generation and characterisation of spheroid cells enriched with CSCs properties from both breast cancer cell lines.

4.3 Results

4.3. 1 Monolayer Culture of Breast Cancer Cell Lines

The two breast cancer cell lines, MCF-7 and MDA-MB-231, available from ATCC were used in this study. Both cell lines were cultured in the same RPMI–supplemented growth medium in T-75 culture flasks incubated at 37 °C under 5% humidified CO_2 atmosphere. The cells were checked microscopically daily to ensure they were healthy and growing as expected. The cells at 70% confluency were observed to be mainly attached to the bottom of the flask and appeared to be in their expected shape and refracting light around their membrane (Figure 4.2). The morphology and other growth properties of both breast cancer cell lines were summarised in Table 4.1. The growth characteristics for both cell lines observed in this study were found to be similar to the characteristics reported in the ATCC product sheet information. Fresh growth medium was added to the cells to replenish the nutrients and to remove the by-products of the cells, as to keep them at an optimal culture condition. The cells were sub-cultured at their log phase before they reached confluence. The cell growth properties, such as population doubling time (PDT) and splitting ratio were determined to ensure the cells were at their optimal condition after post-freeze recovery. PDT was determined to be approximately 38.9 hours and 32.4 hours for MCF-7 and MDA-MB-231 breast cancer cell lines respectively (Figure 4.3). Since MDA-MB-231 cells had a higher doubling time compared to MCF-7 cells, MDA-MB-231 cells were split at 1:6 ratio as compared to 1:4 ratio for MCF-7. The passage number which is the number of sub-cultures of the cells was recorded. Earlier passage cells (less

than 10 passages) were cryopreserved as a backup supply and to ensure that all the cells used throughout the study were at a certain passage number to minimize any possible variations that might occurred.

4.3.2 Mycoplasma-free Cultures

Mycoplasma-specific PCR was performed to check any mycoplasma contamination in the breast cancer cell lines cultures. Prior to that, DNA from the two breast cancer cell lines were extracted at passage 5 and 10, and their concentration, purity, and integrity of the extracted DNA were determined to ensure that the DNAs were sufficient and qualified to be used for the PCR. All the extracted DNAs showed high purities of A260/A280 ratios of more than 1.8, concentration of more than 50 ng/µL (Table 4.2) and no degradation (intact bands) (Figure 4.4). The ratios of A260/A230 for all extracted DNA were 2.0 or higher indicating the absent phenol residuals or other contaminants from the DNA extraction process. In order to verify the cultured MCF-7 and MDA-MB-231 in the laboratory were free of mycoplasma contaminations, mycoplasma-specific PCR reactions were performed on the two different passages of the cells, namely at Passage 5 (P5) and Passage 10 (P10). Cultured cells which are positive for mycoplasma are easily recognised by a distinct PCR product ranging from 502 to 520 bp on the agarose gel. All the cultured breast cancer cell lines were free of mycoplasma contamination as no mycoplasma-specific band appeared in the gel (Figure 4.5). No contamination of PCR reagents was detected as no band was present in the negative control

water control. Hence, the cells were safe to be used for subsequent experiments.

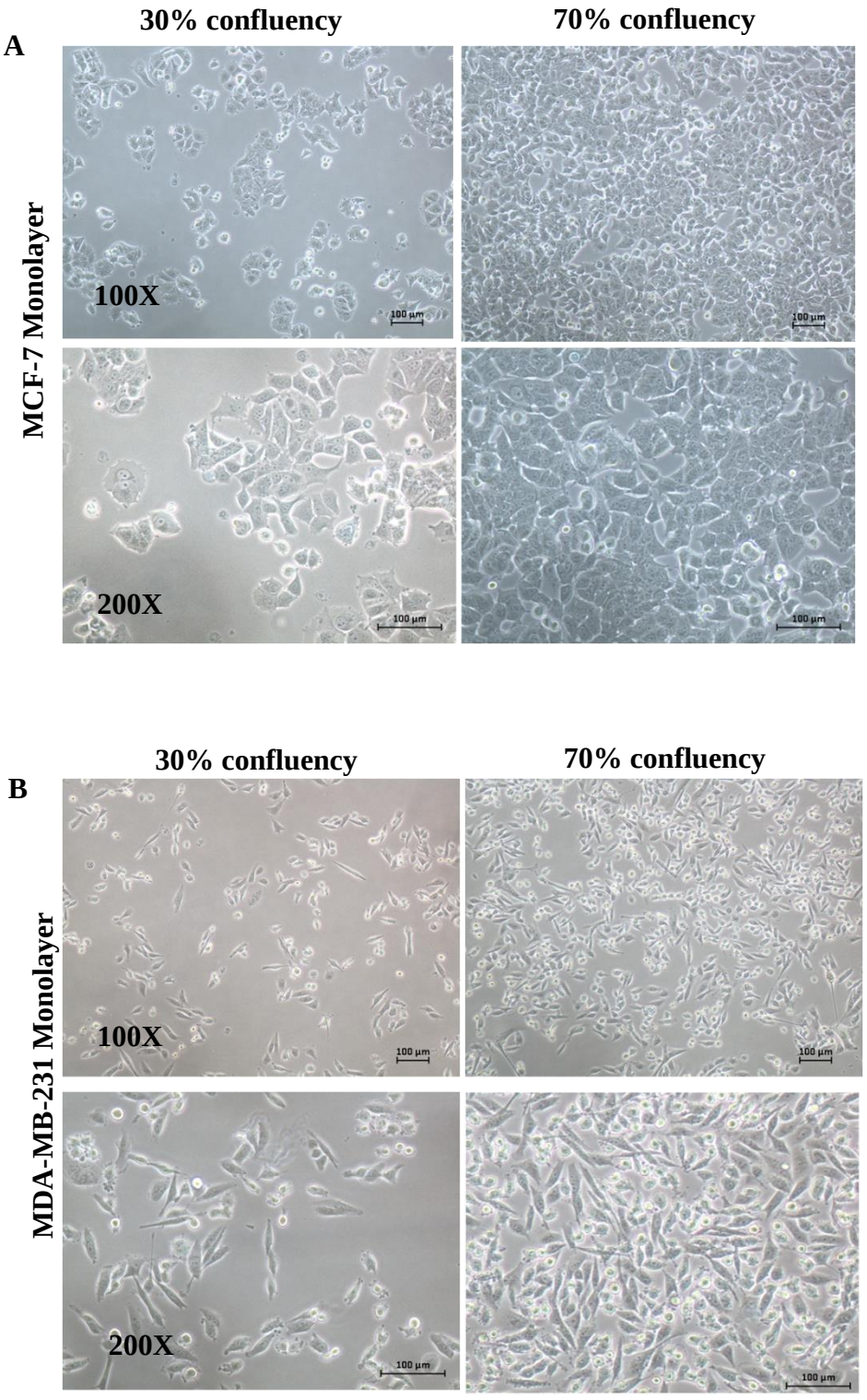


Figure 4.2. Morphology of the breast cancer cell lines at different confluency. Images were at 100X and 200X magnifications.

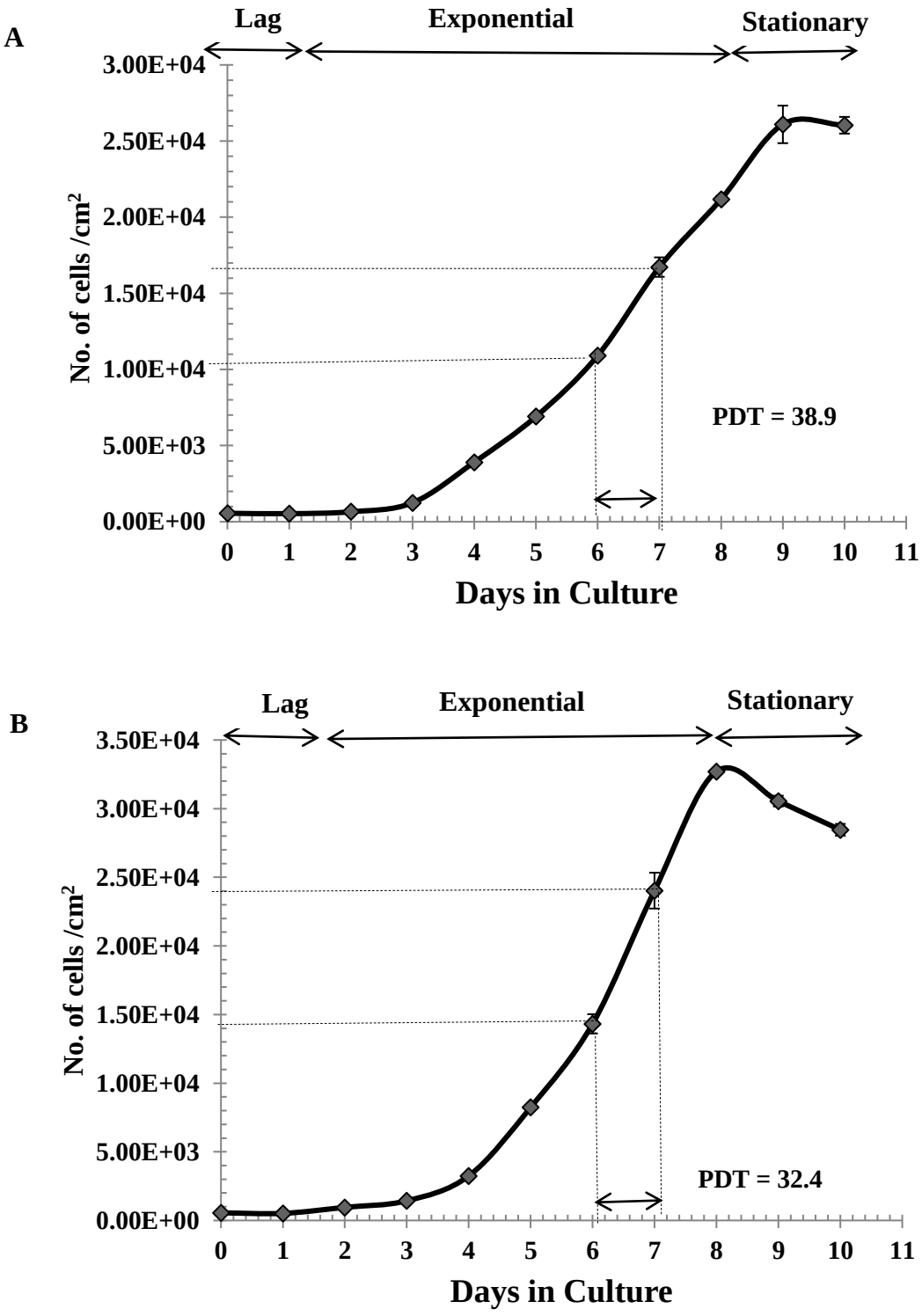


Figure 4.3. Growth curve for MCF-7 and MDA-MB-231 monolayer cultures. Population doubling times (PDT) were determined from the growth curve as indicated in the graph above. (A) MCF-7 monolayer and (B) MDA-MB-231 monolayer.

Table 4.1. Characteristics of the MCF-7 and MDA-MB-231 breast cancer cell lines used in this study.

Characteristics	MCF-7	MDA-MB-231
Shape	Epithelial-like of adherent polygonal shape	Epithelial-like of spindle or stellate shape
Tissue origins	Pleural effusion, metastatic site*	Pleural effusion, metastatic site*
Growth pattern	Slow-growing cells**	Fast-growing cells**
Oestrogen receptor expression	Yes*	No*
Time to reach 70% confluency	5 days*, 6 days**	5.5 days **
Population doubling time	38 hours*/38.9 hours **	38 hours*/32.4 hours **
Splitting ratio	1:5*/1:4**	1:4 to 1:5*/1:6**
Number of cells/cm^2at 70% confluence	2.0 x 10^4 cells/cm^2**	2.5 x 10^4 cells/cm^2**
Size of single cell	13.62 µm**	19.69 µm**
Approximate percentage of viability for each sub-culture	91.73 ± 1.61 %**	92.32 ± 2.39 %**

*Adapted from ATCC webpage http://www.atcc.org/products/all/HTB-22.aspx#generalinformation and http://www.atcc.org/products/all/HTB-26.aspx#generalinformation

**Derived from this study

Table 4.2. Concentration and purity of the extracted DNA samples from the cultured breast cancer cell lines

Cell lines	A260/A280	A260/A230	Concentration (ng/µL)	Yield (µg)
MCF-7 (P5)	1.815	2.217	123.00	3.69
MCF-7 (P10)	1.821	2.155	116.00	3.48
MDA-MB-231 (P5)	1.818	2.143	150.00	4.50
MDA-MB-231 (P10)	1.803	2.142	120.00	3.60

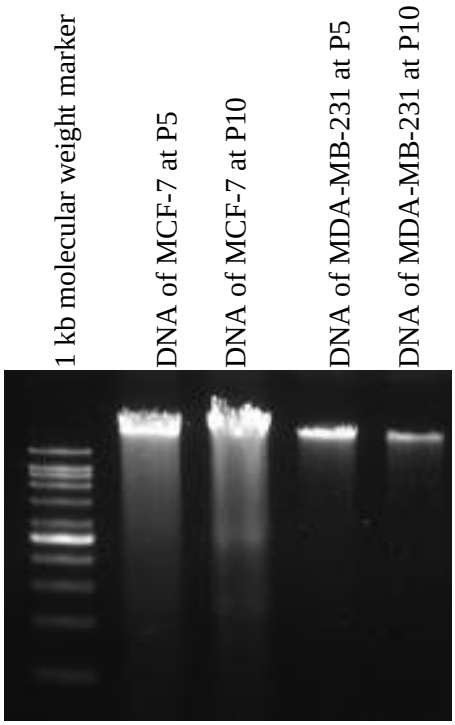


Figure 4.4. Assessment of DNA integrity by agarose gel electrophoresis. Five µL of 50 ng/µL of the extracted DNA from MCF-7 and MDA-MB-231 cells were subjected to 1% agarose gel electrophoresis, stained with gel red, and visualized using UV bioimaging system. P5: Passage 5, P10: Passage 10.

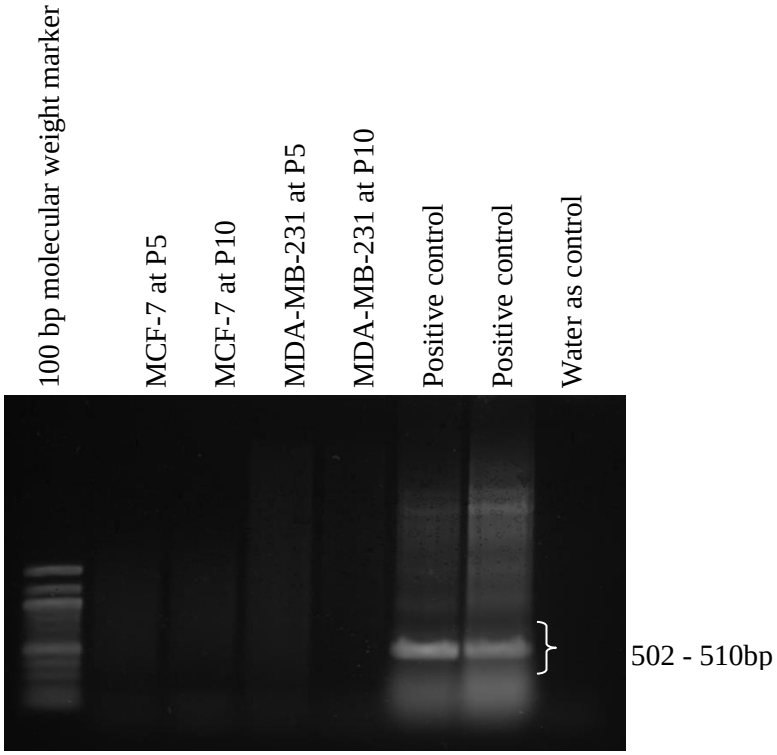


Figure 4.5. The PCR analysis of mycoplasma status in MCF-7 and MDA-MB-231 cultured cell lines. Shown is the 1.0% gel red–stained gel containing the reaction products of the mycoplasma-PCR amplification. P5: Passage 5, P10: Passage 10.

4.3.3 Morphologic Characterisation of Spheroid Cells

The spheroids from both breast cancer cell lines, MCF-7 and MDA-MB-231 were generated using the liquid overlay technique. The 1.0% agar formed a concave surface when polymerised in the bottom of the 96-well plates. MCF-7 and MDA-MB-231 were seeded into the agar-coated plates and were centrifuged to facilitate the cell sedimentation to the bottom of the wells. Cellular aggregation promoted the formation of a single centrally located spheroid of reproducible size. The whole process for spheroid generation is illustrated in Figure 4.6. Under serum-free condition, the aggregated cells grew into a three-dimensional spheroid within 24-48 hours, and their formation were captured using phase contrast bright field microscope (Figure 4.7). After 96 hours, the homogeneous spheroids were considered to be mature and were easily transferred using Pasteur pipette without any disaggregation. Even when using the same spheroid generation technique, the spheroids derived from MDA-MB-231 cell lines were slightly smaller in size measured by the spheroid diameter at day 4. MCF-7 and MDA-MB-231 spheroids were 623.54 μm and 576.43 μm respectively. The number of cells from the dissociated spheroids in both cell types was reduced as compared to their initial seeding number of 5.00 x 10^4 cells. MCF-7 spheroid was estimated to be 4.80 ± 0.25 x 10^4 cells whereas MDA-MB-231 spheroid was 4.76 ± 0.08 x 10^4 cells. To evaluate the surface morphology of the spheroids, scanning electron microscope (SEM) was performed. SEM micrographs at lower magnification (Figure 4.8) revealed aggregation of thousands of cells that were compacted together to give rise to the spheroidal structure. In both spheroids, cell-cell junctions were visible

through the connected cells to preserve the spheroid architecture. Spheroid MDA-MB-231 had a lesser compact spheroidal structure as compared to MCF-7 spheroid based on their higher porosity. Present of pores within the interior of the spheroid structure allowed the diffusion of nutrients and oxygen to the inner layer of the cells.

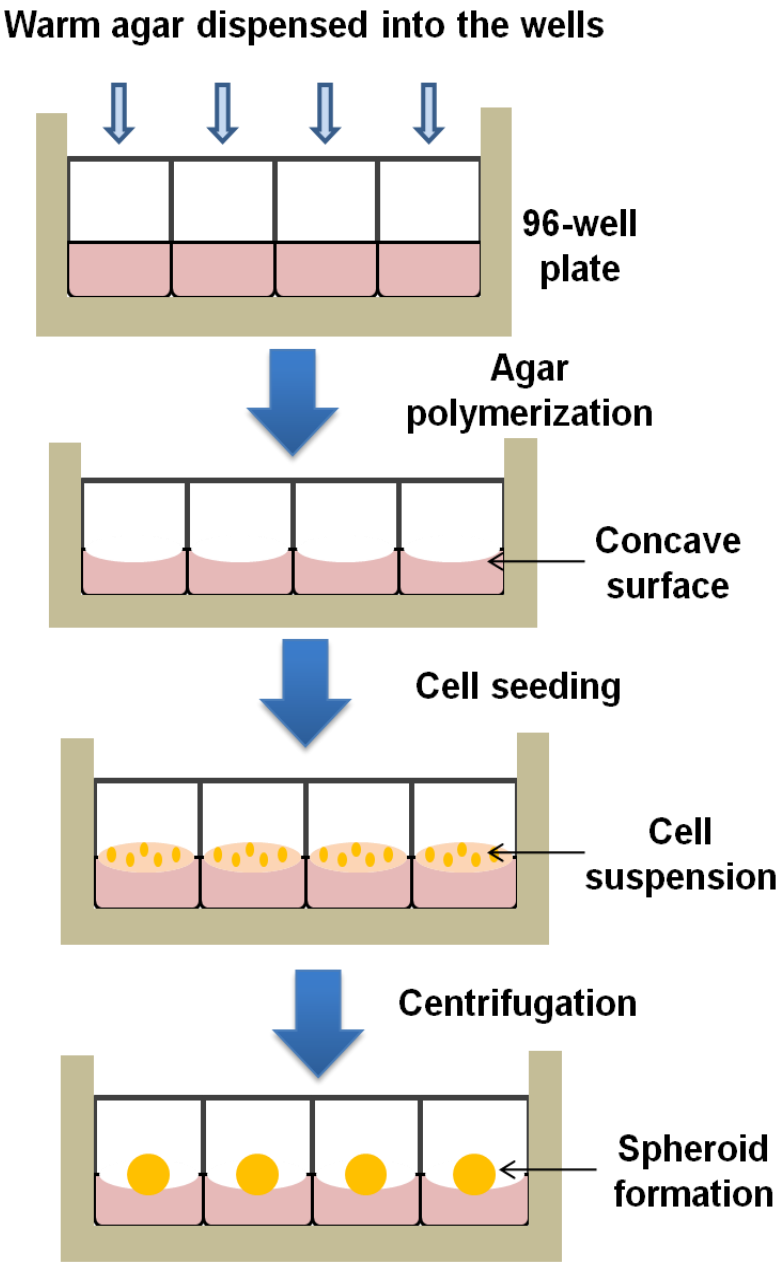


Figure 4.6. Schematic illustration of the agar overlay technique used in spheroid formation. The bottom of the 96-well plate was coated with agar, followed by cell seeding, and centrifugation to induce cell aggregation. Cell suspension was mixed with serum-free medium before dispensed in the 96-well coated plate.

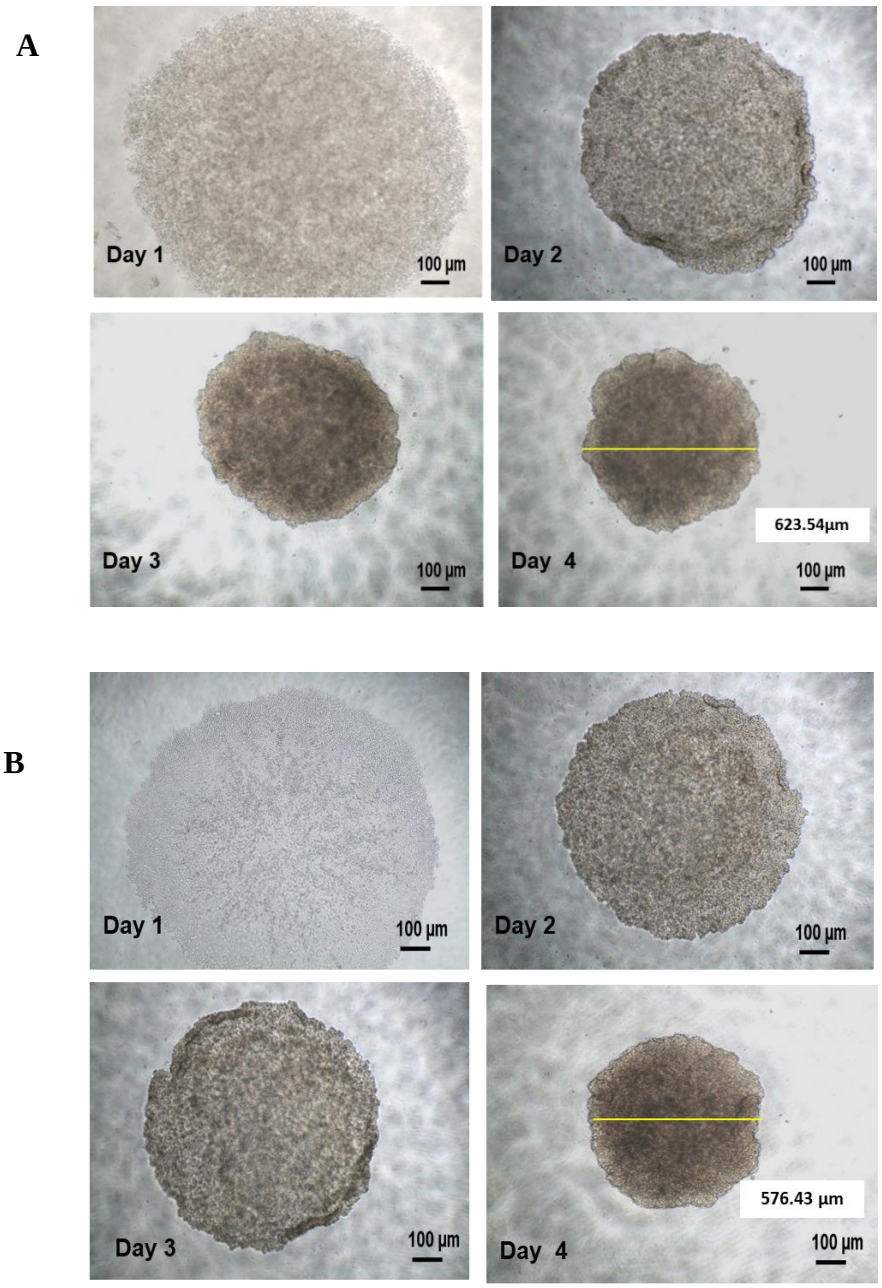


Figure 4.7. The morphology of the spheroid cells from day 1 to day 4. (A) Spheroid derived from MCF-7 breast cancer cell line, (B) Spheroid derived from MDA-MB-231 breast cancer cell line. Both spheroids became more compact as shown by the reduction in size. Images were at 40X magnification.

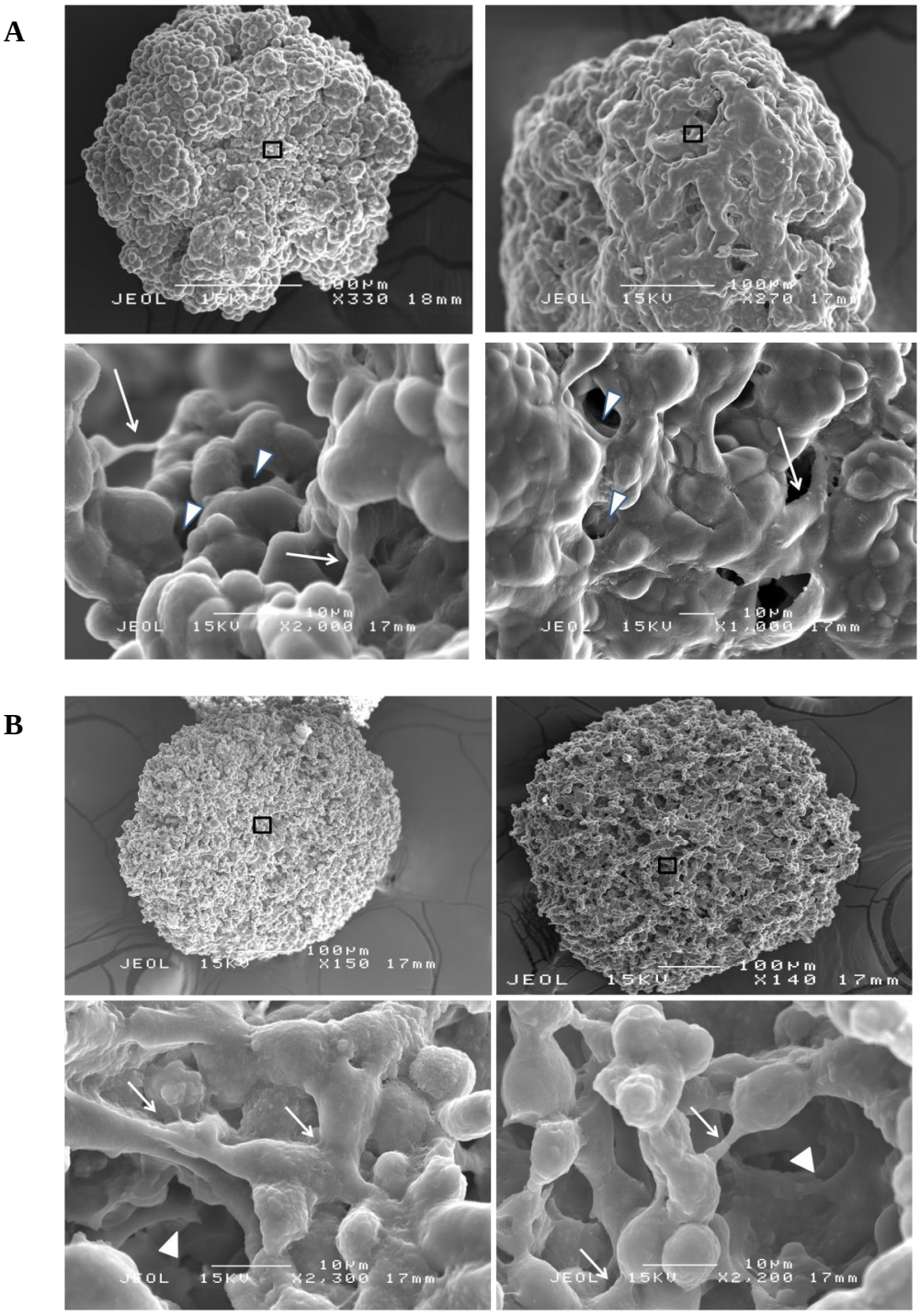


Figure 4.8. The surface morphology of the spheroid cells using Scanning Electron Microscopy. (A) Spheroid derived from MCF-7 breast cancer cell line, (B) Spheroid derived from MDA-MB-231 breast cancer cell line. Both spheroids were intact, but with different pore sizes (arrowheads). Cell-cell junctions (arrows) were visible surrounding the cells. Images were at various magnifications as indicated in the electron micrographs.

4.3.4 Self-renewable Capability of the Spheroid Cells

To further characterise the spheroid cells, the mature spheroid of the two breast cancer cell lines, MCF-7 and MDA-MB-231, were harvested and dissociated into single cells at lower cell density of 200 cells/well and at single cell dilution assay. The single cells were able to be clonally expanded by forming new spheroids from single cell dilution per well as illustrated in Figure 4.9. Dissociated low density cells from both cell lines were also able to form non-adherent, three-dimensional spheroids in serum-free medium under non-adherent condition. The spheroids increased in size from day 1 to day 14, with MDA-MB-231 secondary spheroids measured to be approximately 380.76 ± 25.48 μm as compared to 303.59 ± 22.57 μm in MCF-7.The secondary spheroids were subsequently cultured up to three passages and the sphere-forming efficiency remained relatively constant (Figure 4.10A). The morphology of the secondary spheroids at three different passages of MCF-7 and MDA-MB-231 breast cancer cell lines was shown in Figure 4.10B. The shape and appearance of the sphere was different for both cell types, as MCF-7 spheres appeared to be more rounded and tightly packed whereas MDA-MB-231 spheres appeared to be looser, less compact and formed aggregates.

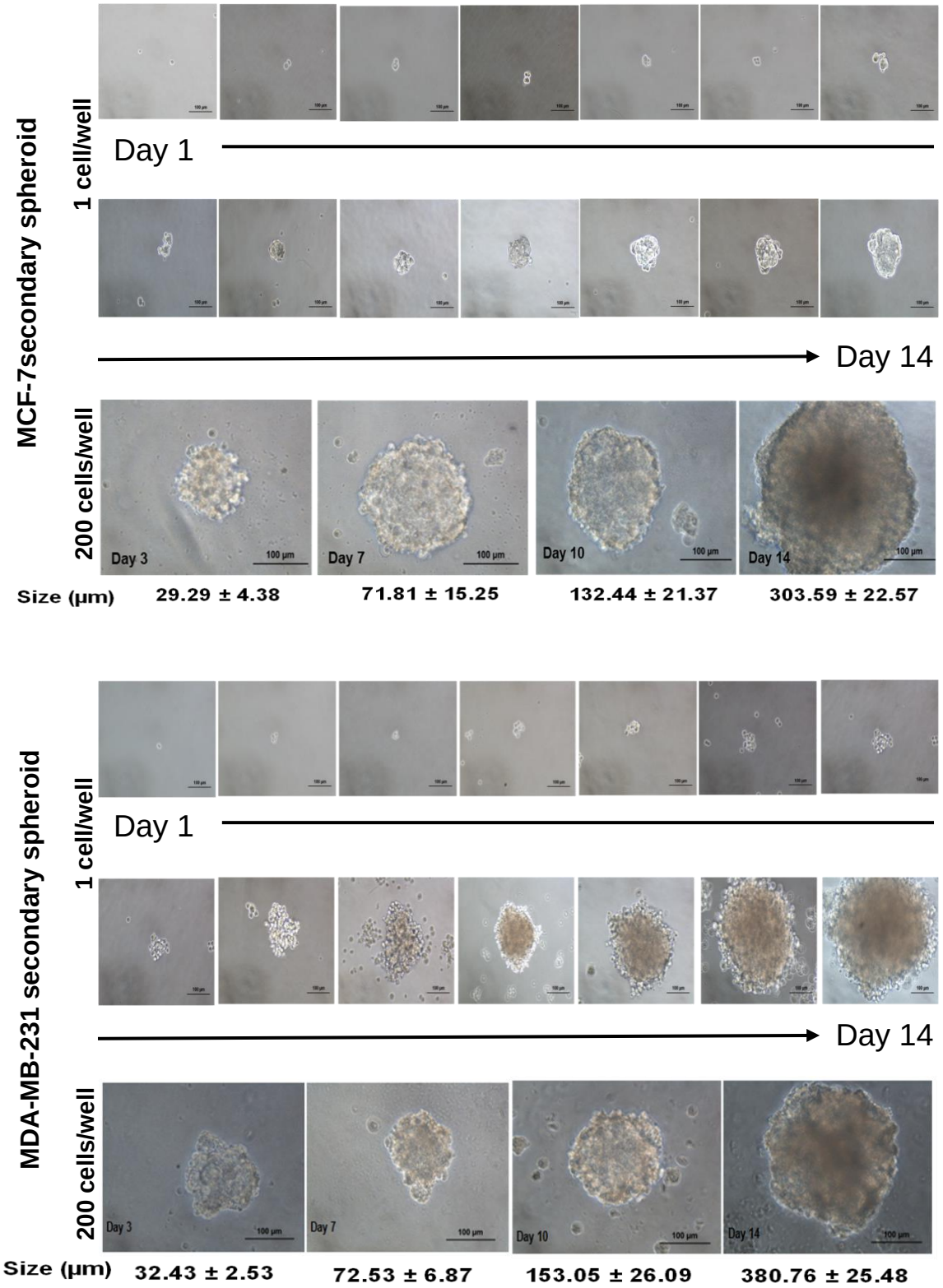


Figure 4.9. Secondary spheroid formation of MCF-7 and MDA-MB-231 spheroid dissociated cells at single cell and 200 cells/well dilution assay. Microscopic images showed the formation of the secondary spheroids derived from a single cell from day 1 to day 14. Progressive increase in size from day 3 to day 14 when the cells were seeded at 200 cells/well. All images were at magnification of 200X.

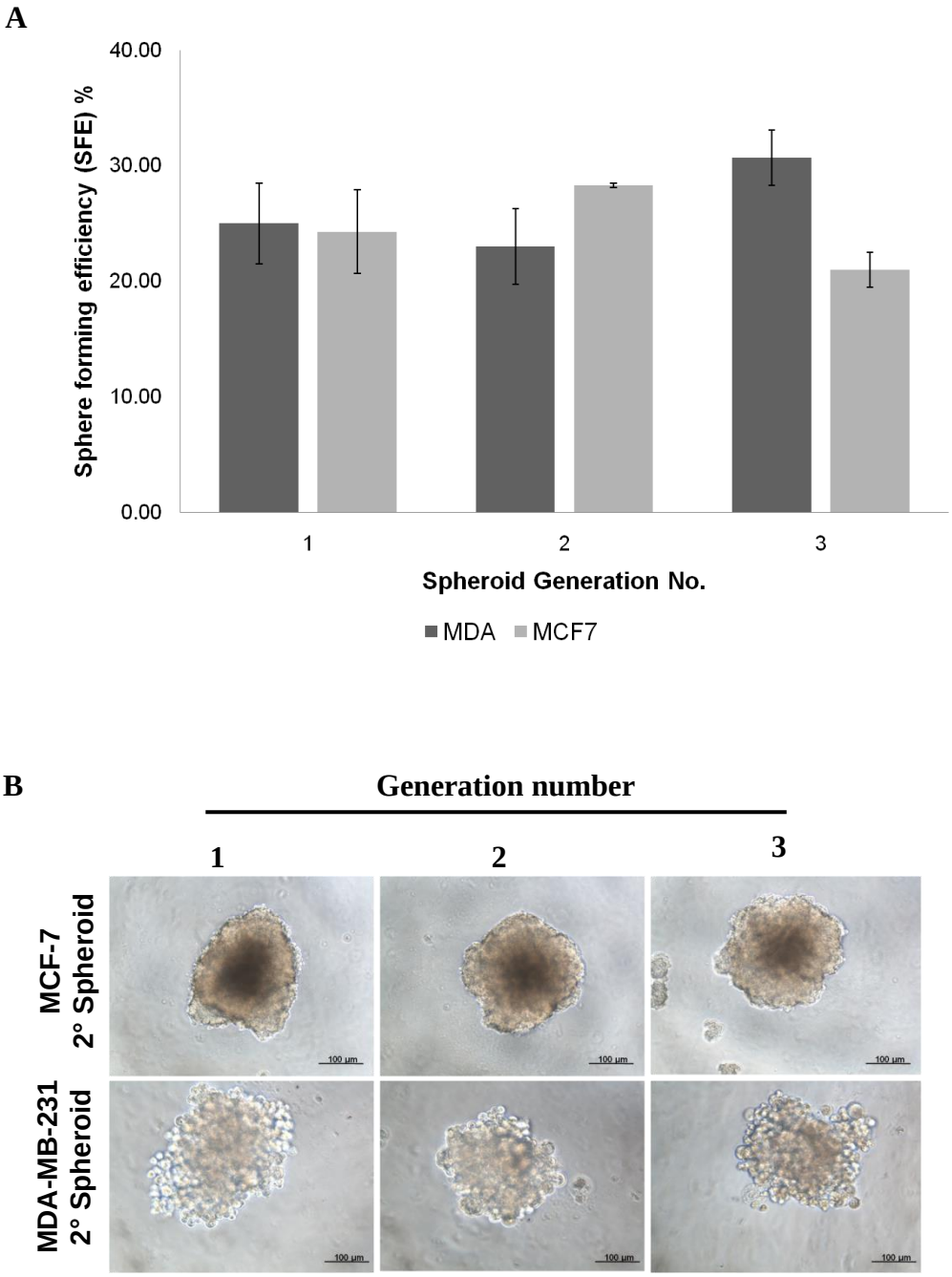


Figure 4.10. Spheroid-forming efficiency (SFE) of secondary spheroids of MCF-7 and MDA-MB-231 spheroid dissociated cells from first to third generation. (A) Bar graph represents the SFE mean calculated by counting the number of spheres formed in a given well and divided by the total number of seeded cells in the well, represented as percentage. Error bars represents standard deviation (SD). Experiments were repeated three times. (B) Representative images of the morphology of the secondary spheroids at different passages of both cell lines. All images were at magnification of 200X.

4.3.5 Immunofluorescence Characterisation of the Spheroid Cells

Expression of stem cells and breast cancer stem cells (CSCs)-related markers were further analysed using immunofluorescence staining. The experiment was performed on the parental monolayer cells which acted as the negative controls and the spheroid cells of both cell lines which were the samples to be investigated. A total of six markers, including three intracellular markers (primary mouse anti-human of Sox2-PE, Nanog-AF, and ALDH1-FITC) and three surface markers (primary mouse anti-human of CD24-PE, CD44-FITC and CD49f-FITC) were used in this study. Parental monolayer cells did not stain for any of the markers except for CD44-FITC. This is expected as CD44 surface marker is a commonly expressed marker for epithelial lineage cells (Chen, Zhao et al., 2018). Spheroid cells were positively stained for all the stem cells and CSCs-related markers except for CD24-PE. Similar staining pattern was observed in secondary spheroids for both cell types. DAPI which stained the nuclei acid blue on all the fixed cells were positive in all the samples. Representative images of the immunofluorescence staining on both breast cancer cell lines were shown in Figure 4.11 to 4.22.

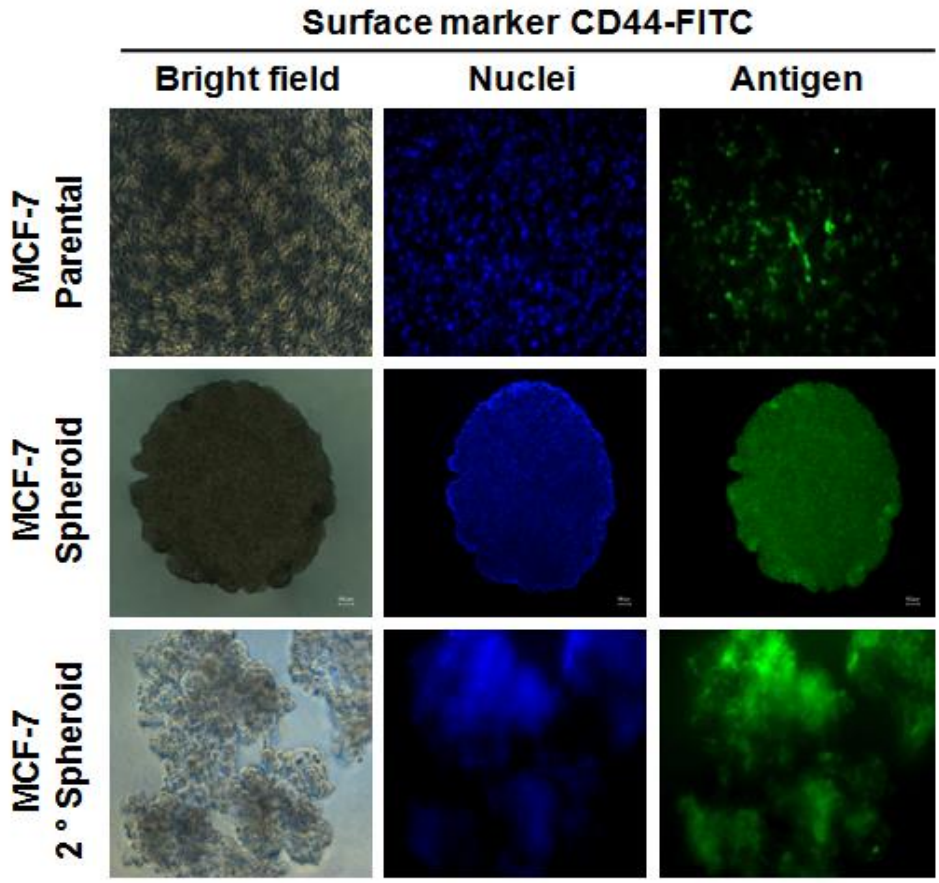


Figure 4.11. Surface staining of CD44 by immunofluorescence staining on MCF-7 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for CD44 surface marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

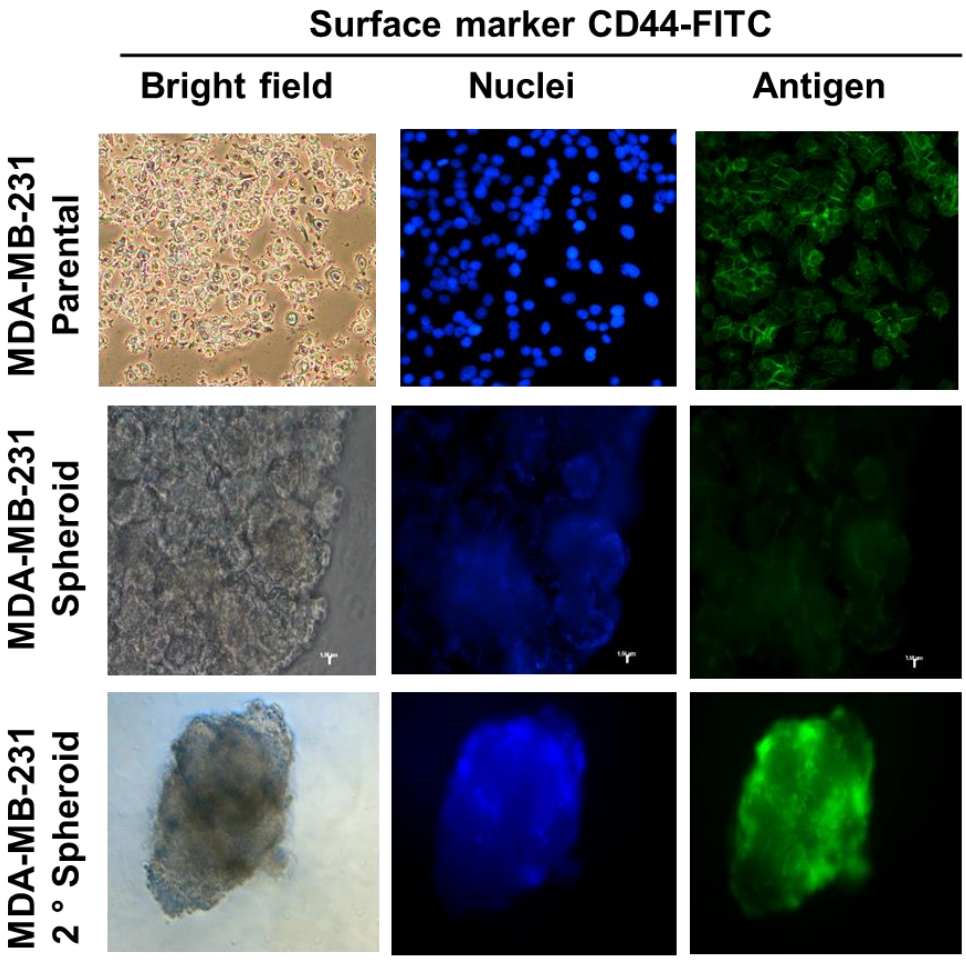


Figure 4.12. Surface staining of CD44 by immunofluorescence staining on MDA-MB-231 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for CD44 surface marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

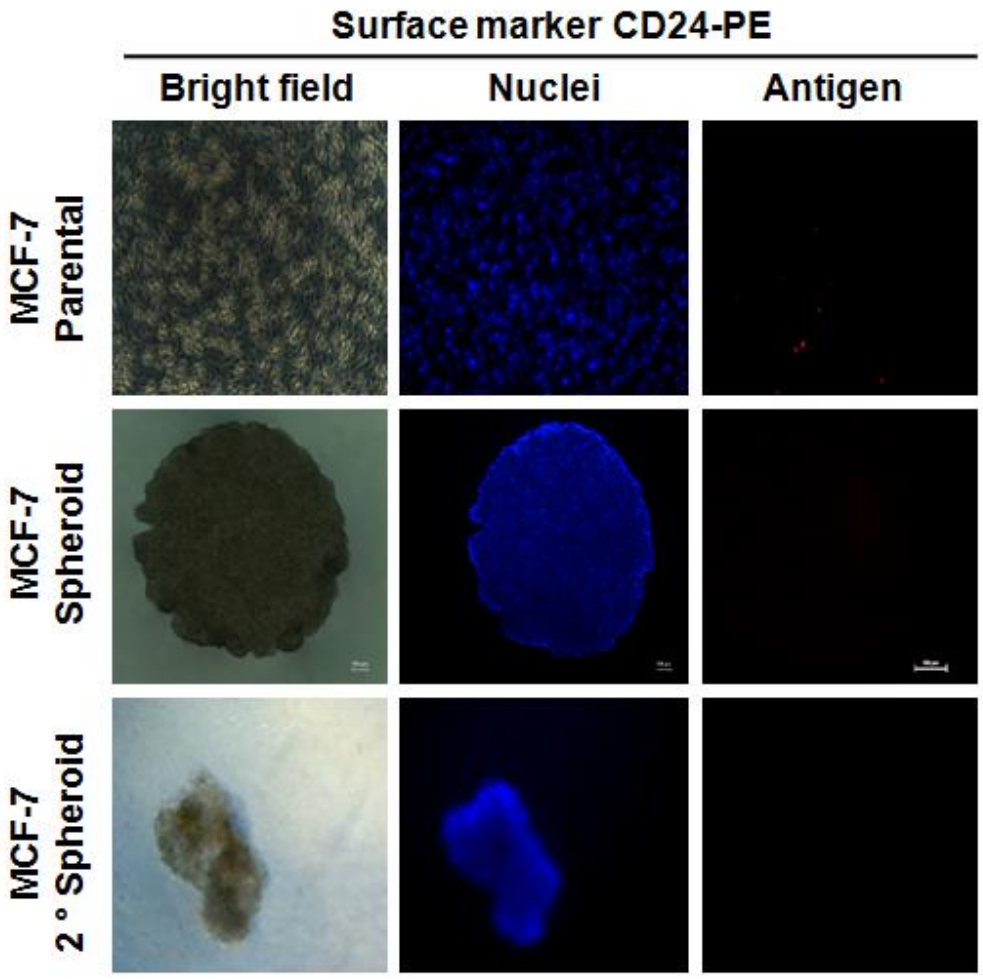


Figure 4.13. Surface staining of CD24 by immunofluorescence staining on MCF-7 spheroids and monolayer cells. Red fluorescence signals were not detected indicated that there were negatively stained for CD24 surface marker in all the samples. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

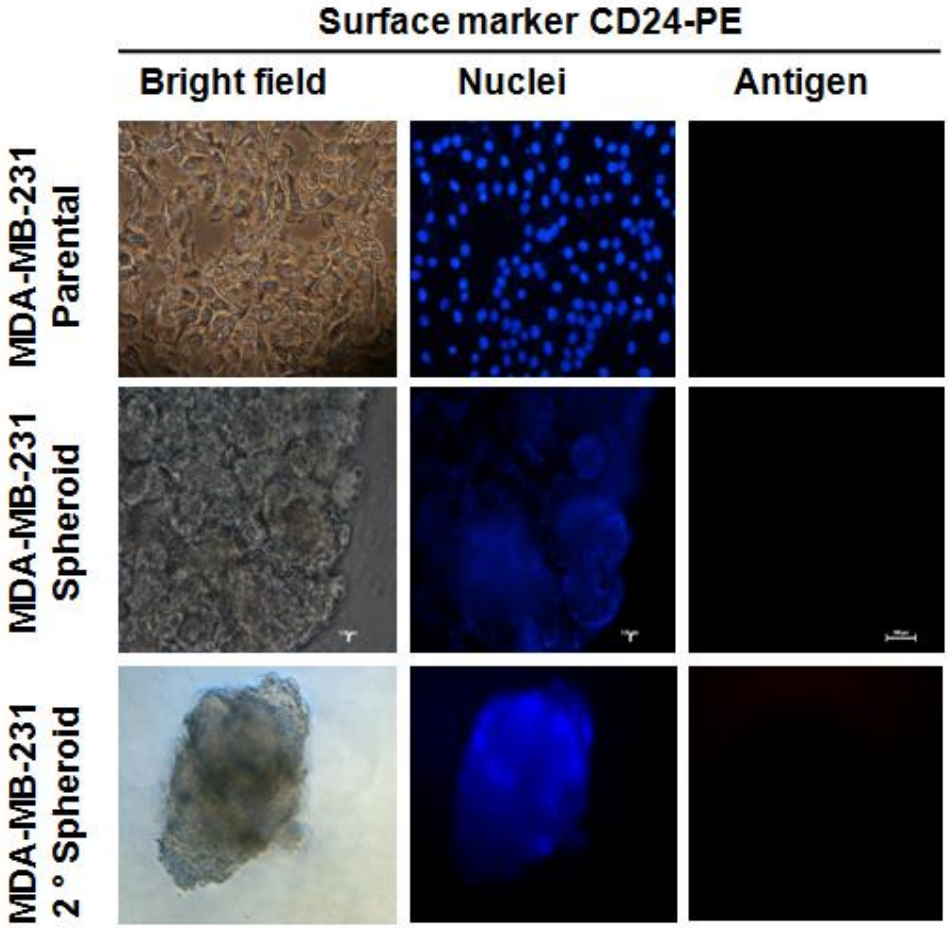


Figure 4.14. Surface staining of CD24 by immunofluorescence staining on MDA-MB-231 spheroids and monolayer cells. Red fluorescence signals were not detected indicated that there were negatively stained for CD24 surface marker in all the samples. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

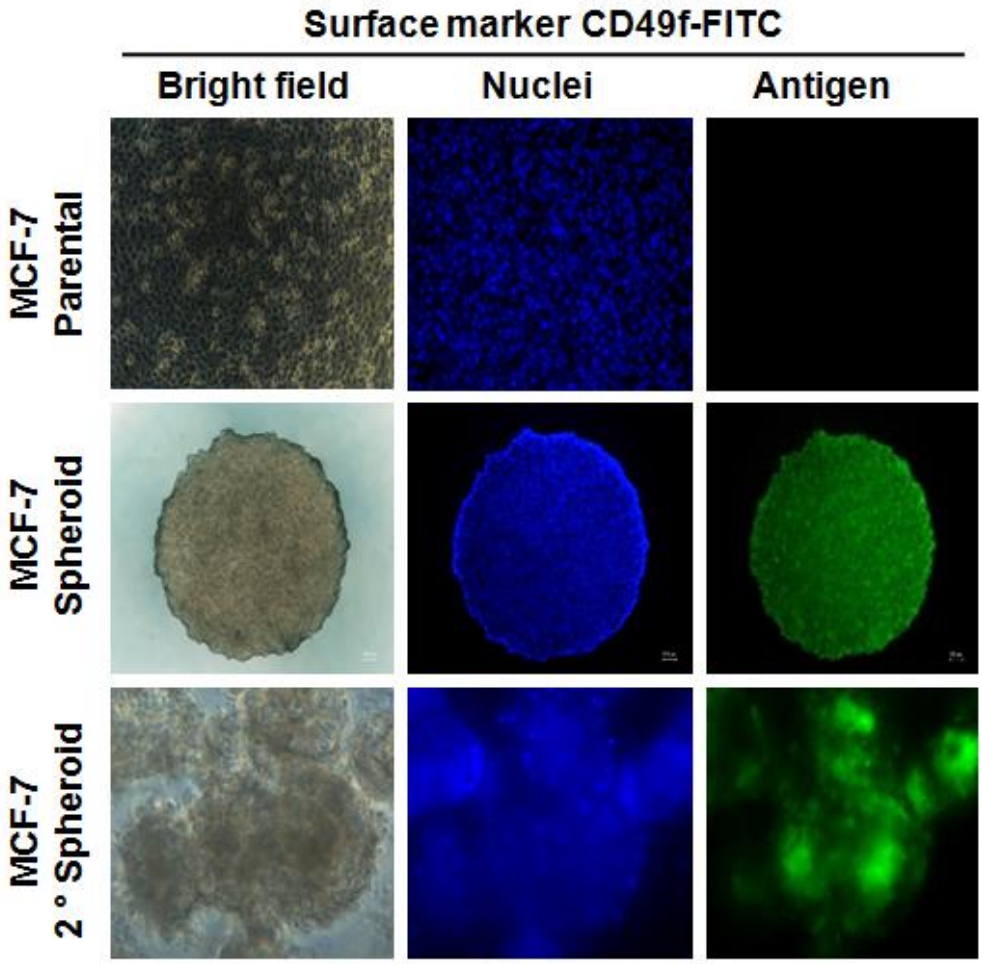


Figure 4.15. Surface staining of CD49f by immunofluorescence staining on MCF-7 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for CD49f surface marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

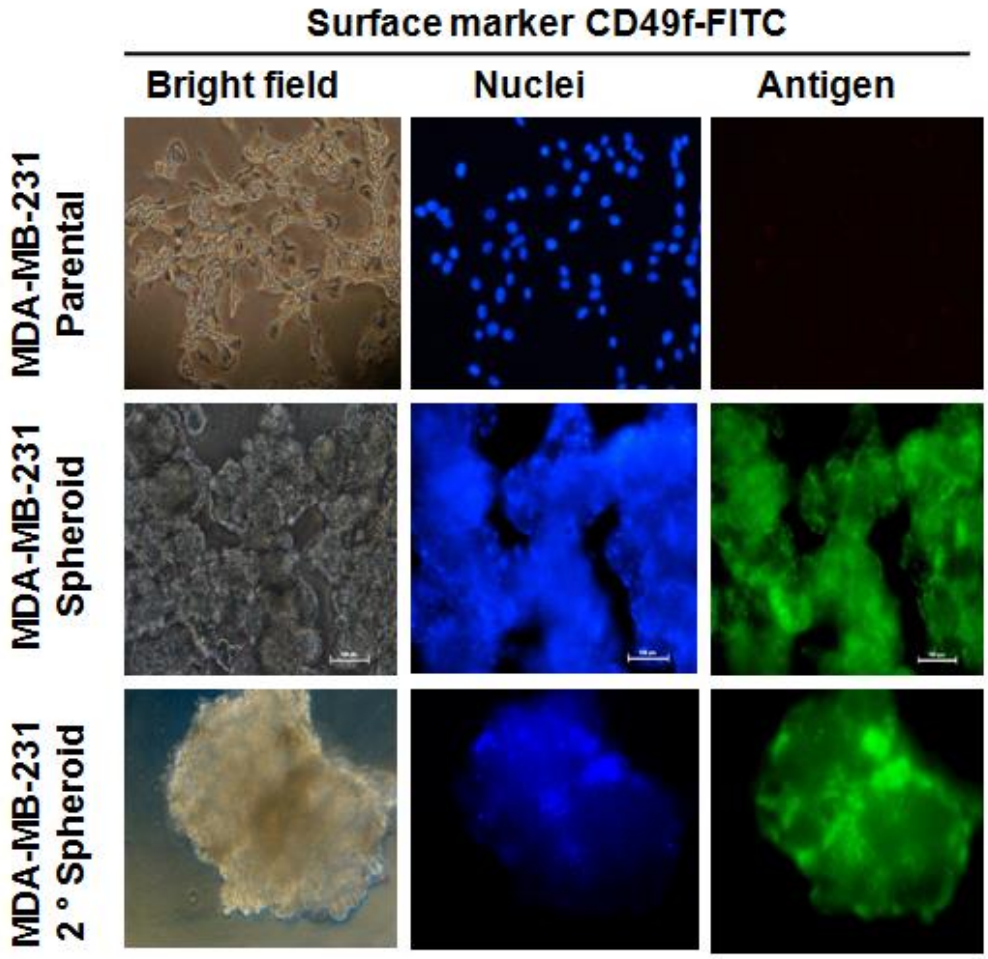


Figure 4.16. Surface staining of CD49f by immunofluorescence staining on MDA-MB-231 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for CD49f surface marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

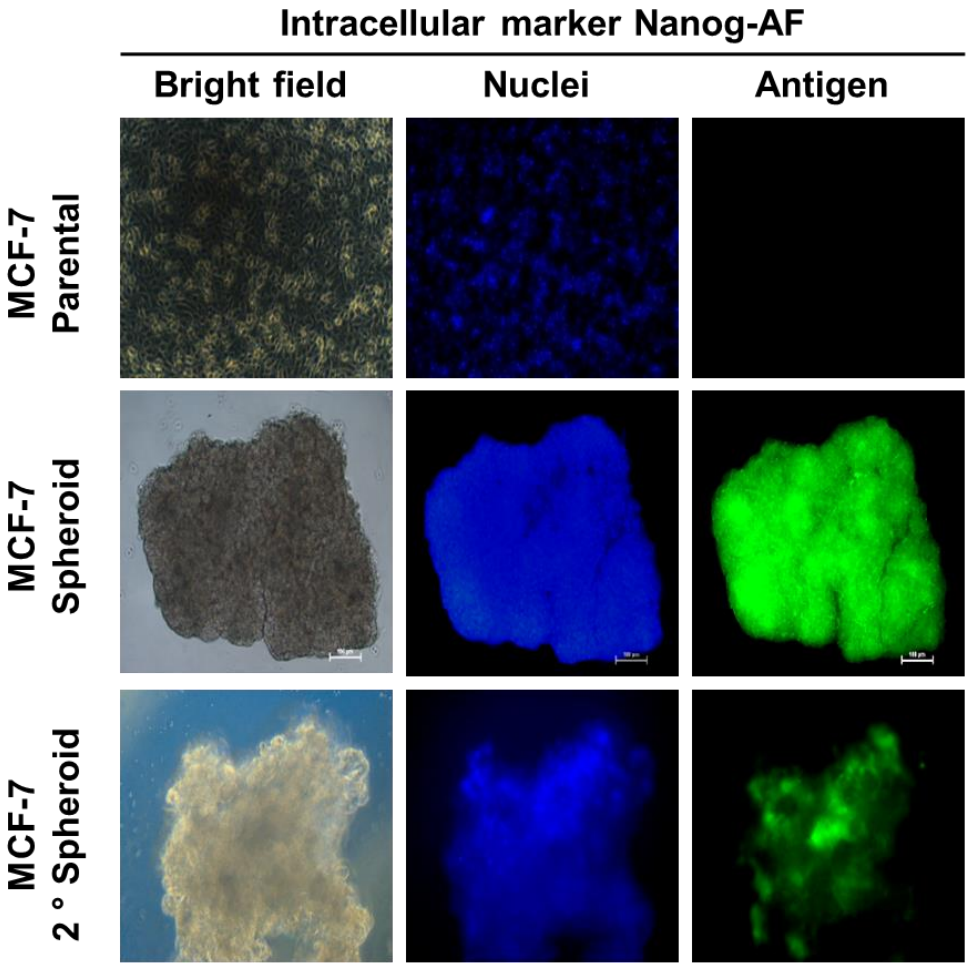


Figure 4.17. Intracellular localization of Nanog by immunofluorescence staining on MCF-7 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for Nanog marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

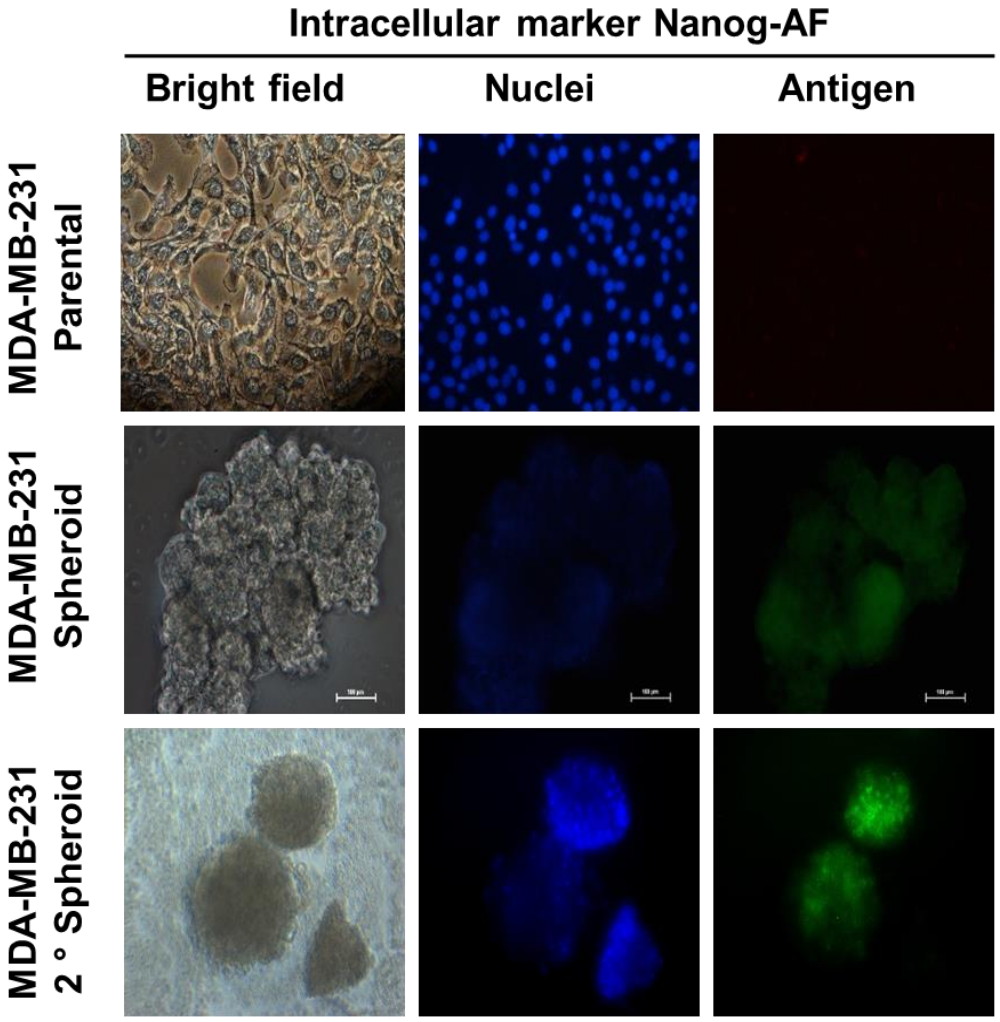


Figure 4.18. Intracellular localization of Nanog by immunofluorescence staining on MDA-MB-231 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for Nanog marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

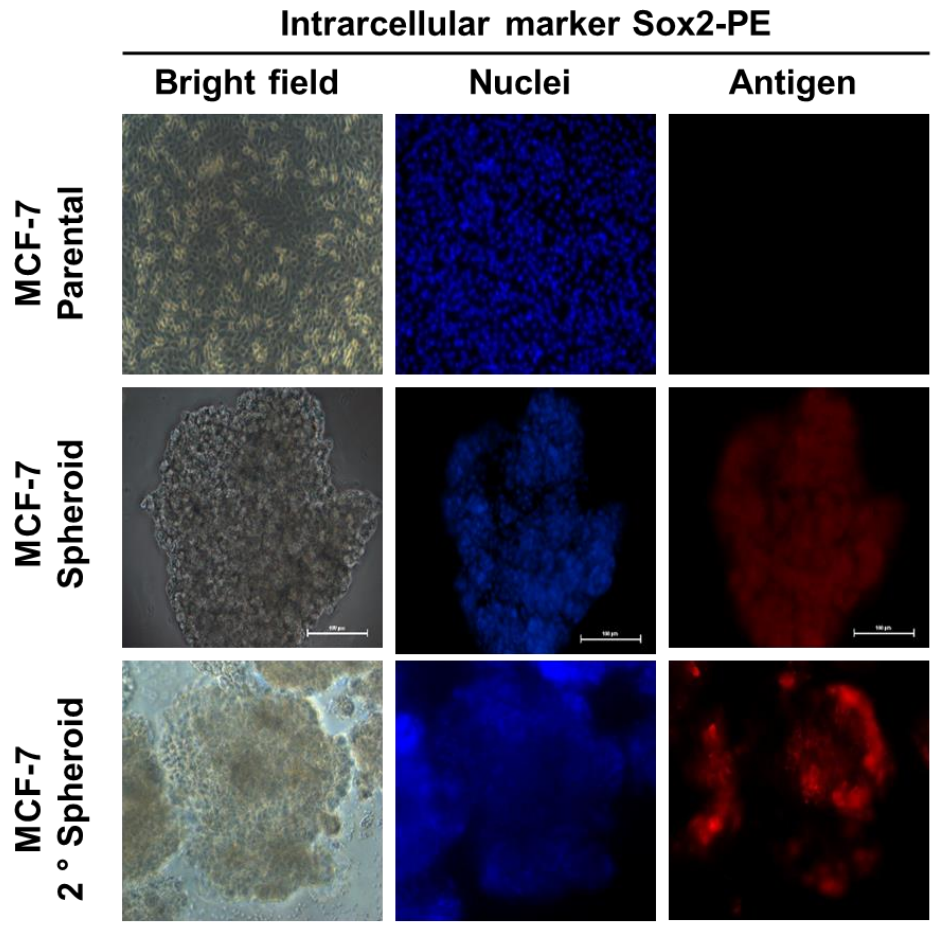


Figure 4.19. Intracellular localization of Sox2 by immunofluorescence staining on MCF-7 spheroids and monolayer cells. Red fluorescence signal indicates positive staining for Sox2 marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

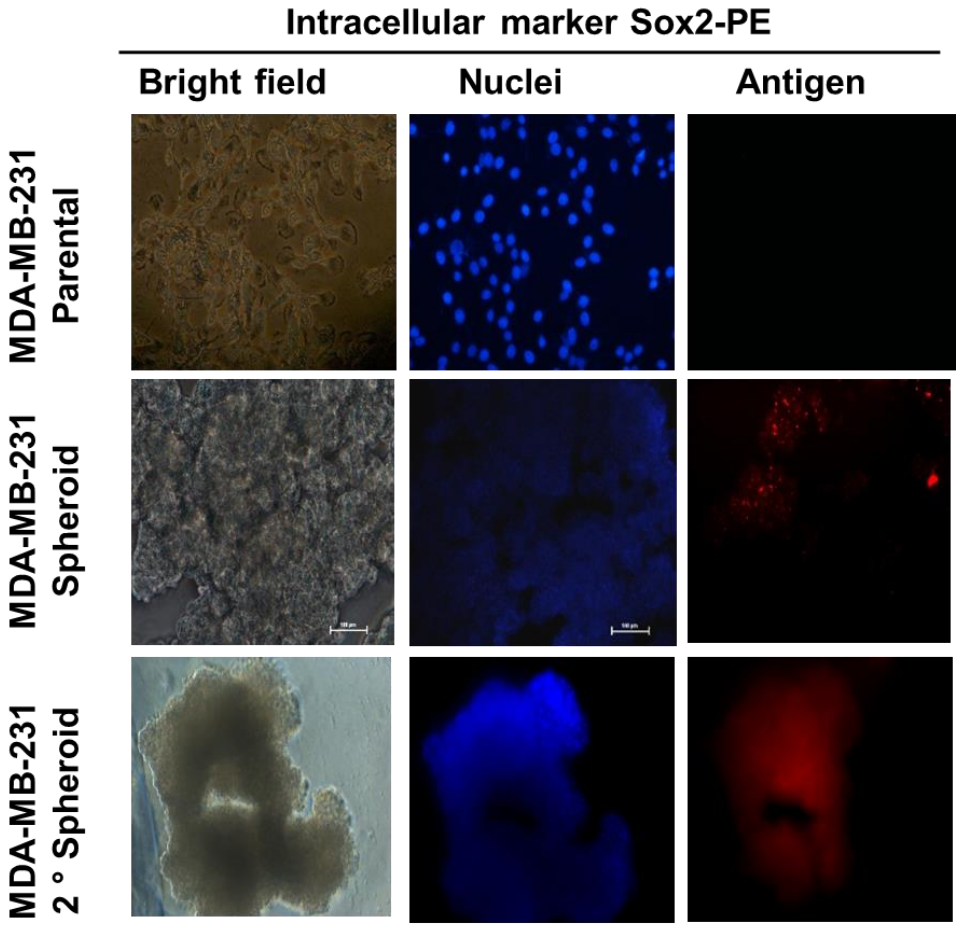


Figure 4.20. Intracellular localization of Sox2 by immunofluorescence staining on MDA-MB-231 spheroids and monolayer cells. Red fluorescence signal indicates positive staining for Sox2 marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

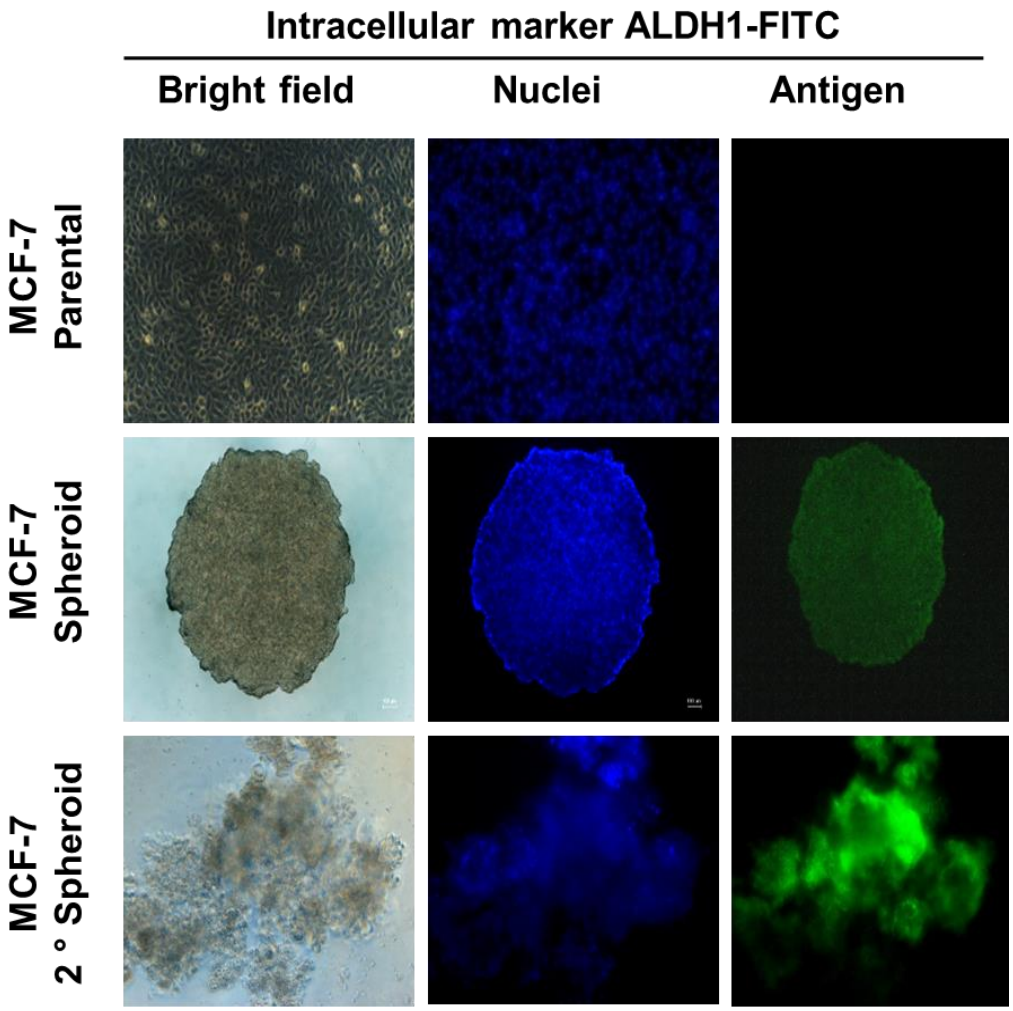


Figure 4.21. Intracellular localization of ALDH1 by immunofluorescence staining on MCF-7 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for ALDH1 marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

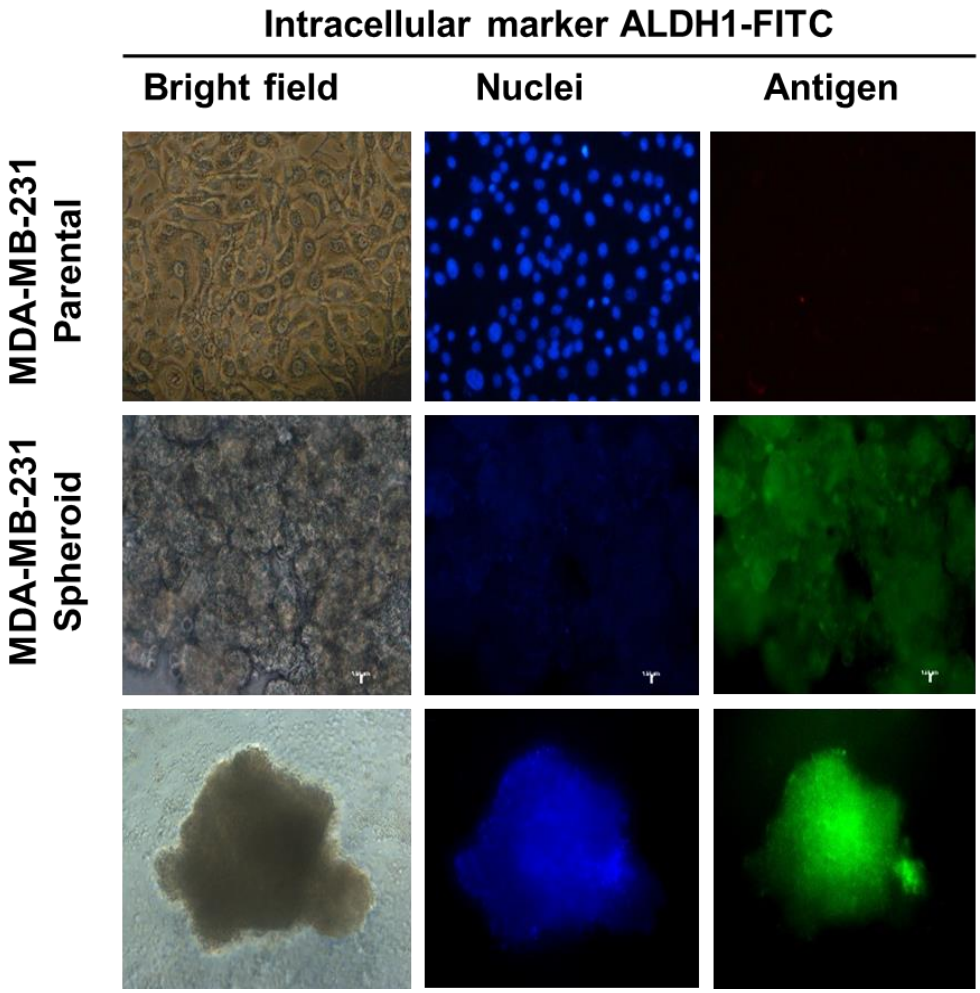


Figure 4.22. Intracellular localization of ALDH1 by immunofluorescence staining on MDA-MB-231 spheroids and monolayer cells. Green fluorescence signal indicates positive staining for ALDH1 marker. DAPI was used for counterstain to indicate the presence of nucleic acid. All images were at magnifications of 40X and 100X.

4.3.6 Enrichment of CD44+/CD24- and ALDH+ Expressions in Spheroid Cells

Flow cytometry analysis was performed to quantitate the expression of breast CSCs surface markers, CD44/CD24 and ALDH activity. Monolayer parental cells were used as the negative controls. The experiment was repeated three times, and the unstained controls and single stained controls were included for each run. The representative dot plots indicating the distribution of cells expressing CD44+/CD24-/low and ALDH+ of the experiments for MCF-7 spheroid cells/parental cells and MDA-MB-231 spheroid cells/parental cells were shown in Figure 4.23 and 4.24. As shown in the figures, the proportion of cells stained positive for CD44 and negative for CD24 were significantly higher in spheroids as compared to its parental counterparts. A total of 18.45 ± 0.51% (mean ± SD; *n*=3) MCF-7 spheroid cells exhibited CD44+/CD24-/low surface markers as compared to 0.63 ± 0.46% (mean ± SD; *n*=3) in parental monolayer MCF-7 cells. Expression of CD44+/CD24-/low in MDA-MB-231 spheroid cells (70.42 ± 2.22%) (mean ± SD; *n*=3) were also found to be significantly higher than its parental counterparts (27.28 ± 1.65%) (mean ± SD; *n*=3).

ALDH-positive cells were 15.38 ± 0.50% (mean ± SD; *n*=3) in MCF-7 spheroid cells which were statistically higher compared to its parental cells which were 1.81 ± 0.70% (mean ± SD; *n*=3). For MDA-MB-231 spheroids, similar trend was also observed where higher ALDH-positive cells population was found to be higher when compared to that of the parental cells (29.43 ±

1.07% compared to 1.64 ± 0.52%) (mean ± SD; *n*=3) (Figure 4.23). The summary of the CD44+/CD24-/low and ALDH+ in spheroid cells of both cell types when compared to their parental counterparts was illustrated in Figure 4.25 and Appendix D.

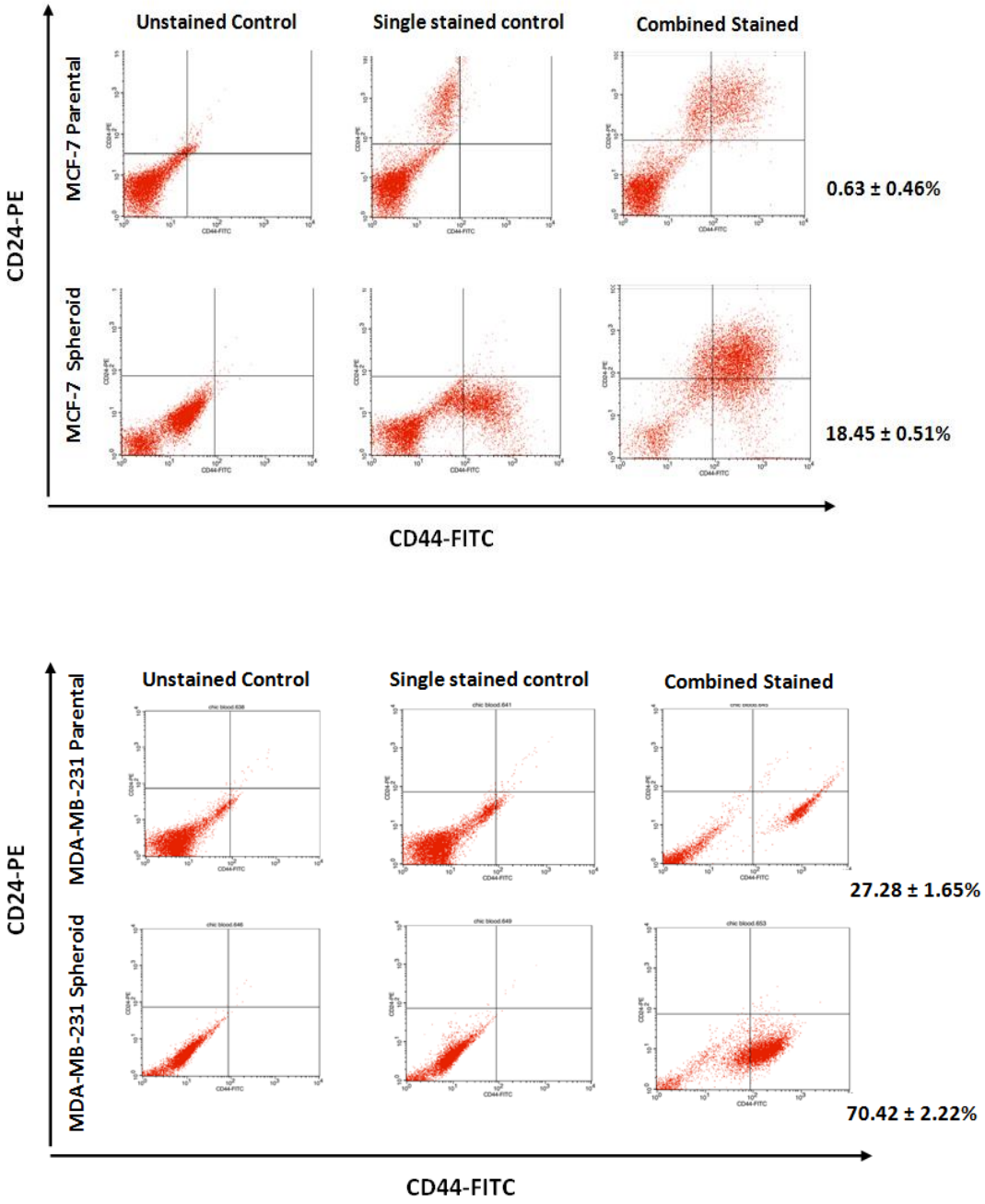


Figure 4.23. Flow cytometry analysis of breast CSCs-related surface markers CD44+/CD24-/low in MCF-7 and MDA-MB-231 spheroid and their parental cells. Representative dot plots showed the expressions of both cell types from three independent experiments. Expression of CD44+/CD24-/low surface markers was found to be increased in spheroids cells compared to its parental.

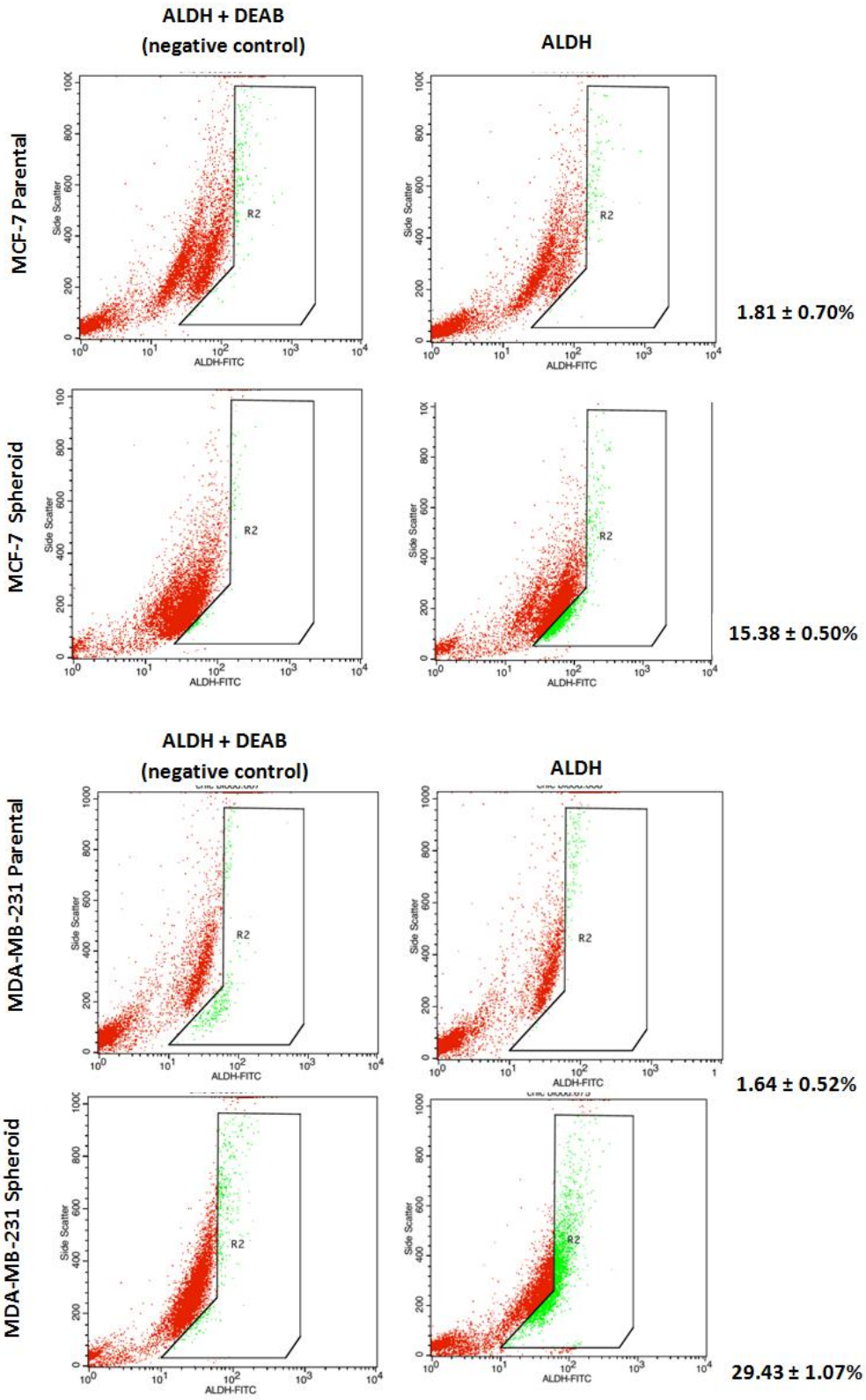


Figure 4.24. Flow cytometry analysis of ALDH activity in MCF-7 and MDA-MB-231 spheroid and their parental cells. Representative dot plots showed the expressions of both cell types from three independent experiments. ALDH activity was expressed higher in spheroids cells compared to its parental. R2 are the region of ALDH-positive cells.

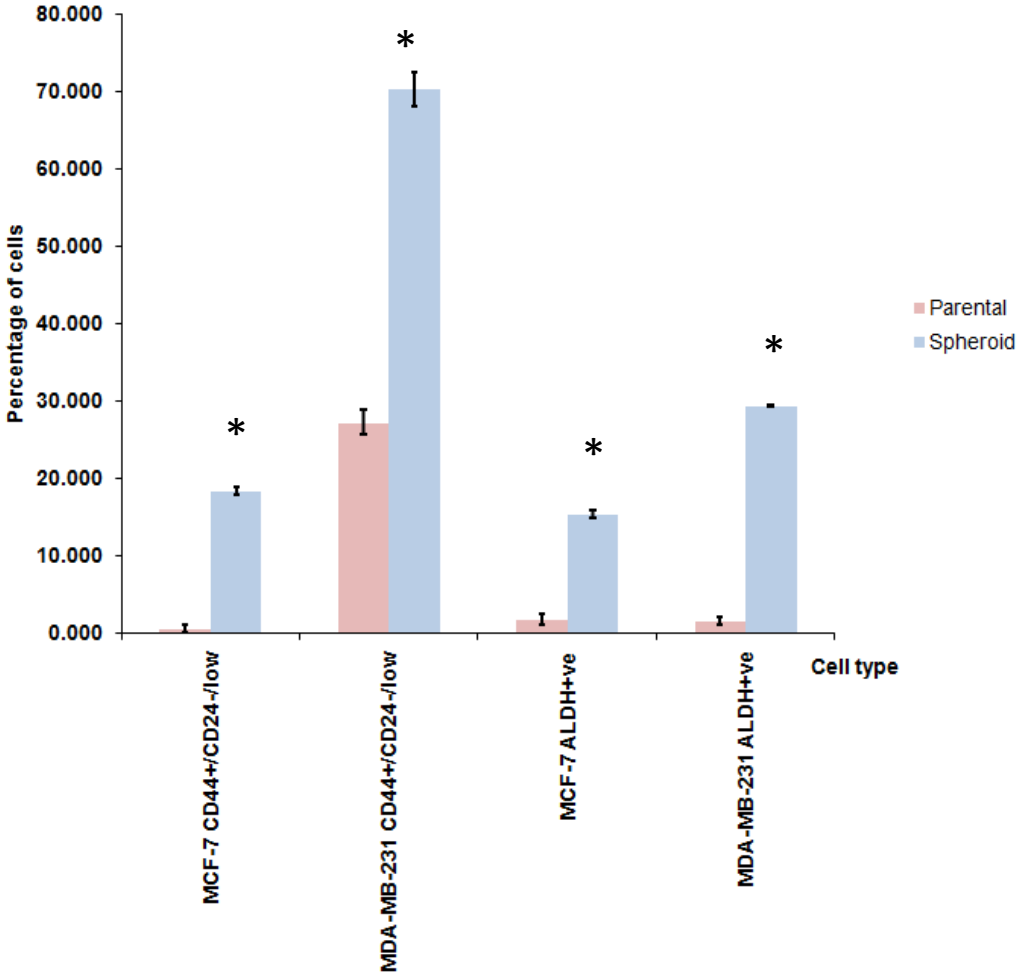


Figure 4.25. Mean expression of the breast CSCs CD44+/CD24-/low markers and ALDH-positive cells activity in MCF-7 and MDA-MB-231 spheroid and their parental cells. Columns, mean; bars, SD; *, statistical significance at p <0.05.

4.3.7 Higher Drug Resistance in Spheroid Cells

Cell titration assays were carried for both MCF-7 and MDA-MB-231 monolayer cell to determine the optimal seeding density prior to drug sensitivity assays. The monolayer cells were diluted into 7 series cell dilution (from low cell number to high cell number) with 100 µL of cells/well, in triplicates, and MTT assays were performed on the next day. As shown in Figure 4.26, the cell seeding number for both breast cancer cell lines were determined to be 5 $x10^3$ cells/well. This cell number was chosen as it yielded an absorbance of not more than 1.25, and the value fall within the linear portion of the regression line. Drug sensitivity assays of the parental versus the spheroids subjected to three chemotherapeutic drugs (tamoxifen, doxorubicin, and cisplatin) were assessed to examine whether the spheroid cells possessed higher chemoresistance ability. For parental cells, the drug treatment and MTT assay were performed according to standard monolayer protocol. For spheroids, drug treatment and MTT assays were performed in three-dimensional (3D) format and also dissociated into single cells before replated and treated as in monolayer cultures, known as two-dimensional (2D) format. The dose-response curves for the MTT assays of the parental and spheroids cells treated with three different drugs were shown in Figure 4.27 and Figure 4.28. Overall, the drug inhibitory concentration (IC_{50}) of the spheroid cultures, regardless in 3D or 2D culture conditions were significantly higher when compared to the parental cells (Table 4.3 and Appendix E & F). For MCF-7 cells, the IC_{50} values under 100 µg serially diluted drugs for 3D spheroids were 2.9-fold, 18.8-fold and 5.4-fold, whereas for 2D spheroid, the IC_{50} values were

2.3-fold, 13.0-fold, and 5.4-fold for tamoxifen, doxorubicin and cisplatin respectively, when compared to parental cells. For MDA-MB-231 cells, the IC_{50} values for 3D spheroids relative to parental cells were 3.67-fold, 11.24-fold and 9.00-fold higher for tamoxifen, doxorubicin and cisplatin. Similarly, spheroid cells in 2D condition of MDA-MB-231 were 3.63-fold, 16.87-fold, and 2.30-fold higher resistance to the same chemotherapeutic drugs compared to their parental counterparts.

Microscopic images performed on spheroids (3D condition) at 96 hours post-treatment at IC_{50} inhibitory drug concentration showed reduction in volume and size of the spheroids (Figure 4.29 and Figure 4.30). Interestingly, when dissociated spheroids replated in monolayer condition and subjected to the same drug treatment, smaller spheroids were found to be emerging from the monolayer cells (Figure 4.31 and Figure 4.32). The morphological changes of the parental MCF-7 and MDA-MB-231 cells induced by the different chemotherapeutic drugs were also examined. Treated monolayer cells showed common morphological features of apoptotic cells such as shrinkage of the cells and cell membrane blebbing (Figure 4.33 and Figure 4.34).

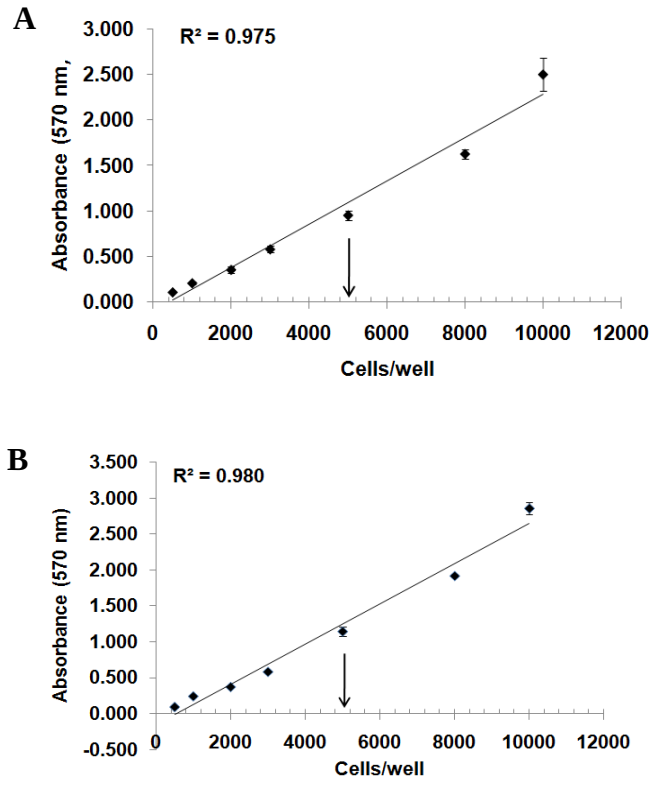


Figure 4.26. Cell titration experiment using MCF-7 and MDA-MB-231 monolayer cells. Arrows indicate the optimal cell number. The regression coefficients, R-squared were indicated on upper left of the graphs. (A) MCF-7 monolayer, (B) MDA-MB-231 monolayer.

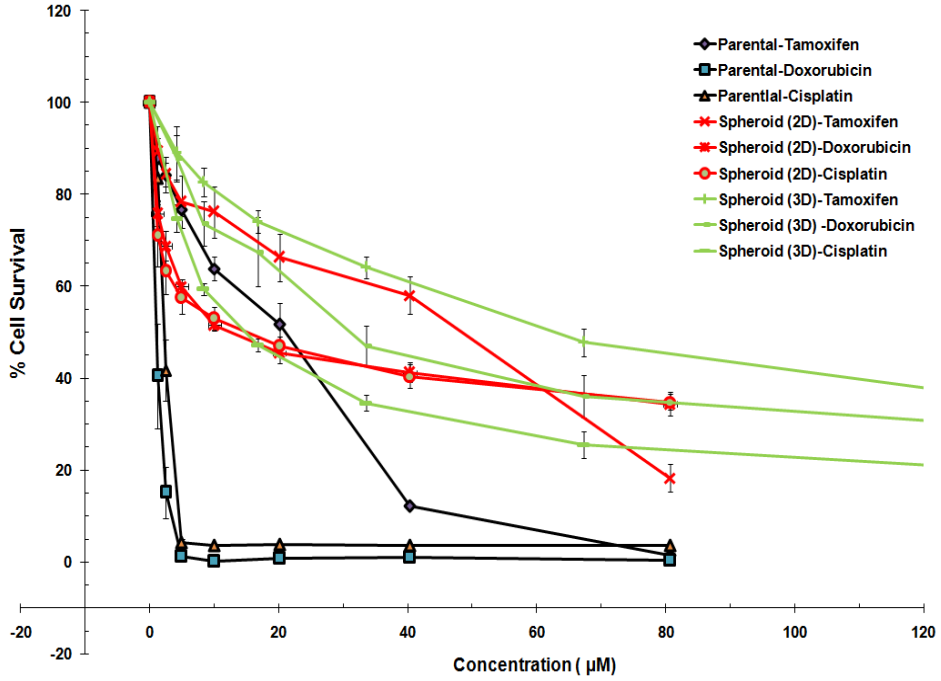


Figure 4.27. MTT assay of MCF-7 parental and spheroid cells treated with three chemotherapeutic drugs. Dose-response survival curves based on MTT calorimetric assay at 96 hours post-treatment. All data are expressed as means ± SD, $p<0.05$ when compared to the parental cells.

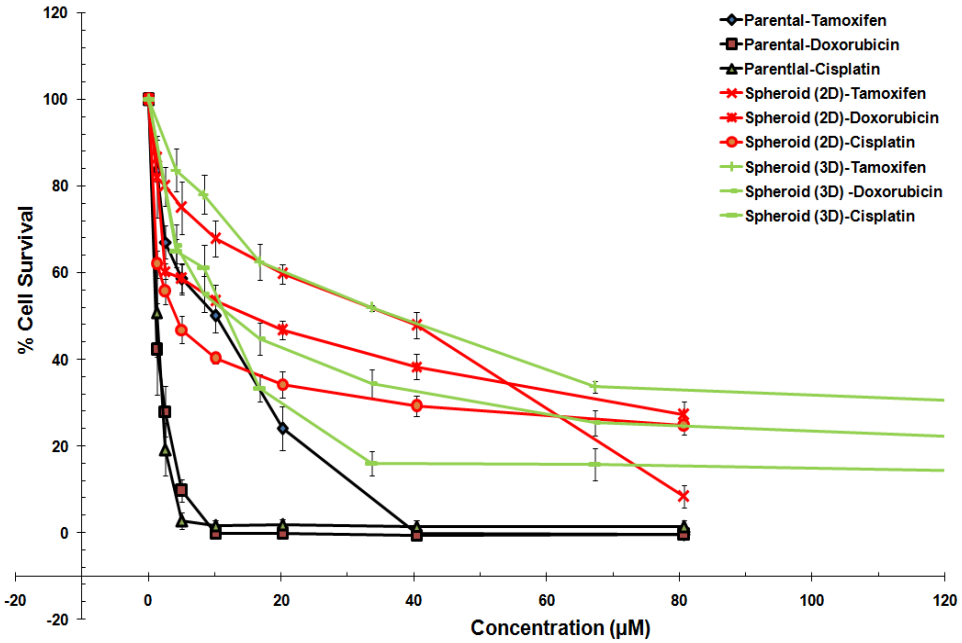


Figure 4.28. MTT assay of MDA-MB-231 parental and spheroid cells treated with three chemotherapeutic drugs. Dose-response survival curves based on MTT calorimetric assay at 96 hours post-treatment. All data are expressed as means ± SD, $p<0.05$ when compared to the parental cells.

Table 4.3. The IC_{50} values of drugs tamoxifen, doxorubicin and cisplatin in parental and spheroid cells (3D and 2D conditions) of MCF-7 and MDA-MC-231 breast cancer cell lines. Cell survival was determined by MTT assay and the IC_{50} values were determined from the dose-response curves. All data are expressed as means ± SD. *p<0.05 when compared with the parental cells.

	Drug inhibitory concentration, IC_{50} (μM)		
	Tamoxifen	**Doxorubicin**	**Cisplatin**
MCF-7 Parental	20.57 ± 2.25	1.00 ± 0.37	2.54 ± 0.44
MCF-7 Spheroid (2D)	47.62 ± 6.83*	13.25 ± 3.69*	13.68 ± 3.92*
MCF-7 Spheroid (3D)	60.73 ± 7.39 *	18.82 ± 3.05*	13.61 ± 3.54*
MDA-MB-231 Parental	10.28 ± 0.270	0.719 ± 0.131	1.74 ± 0.383
MDA-MB-231 Spheroid (2D)	37.34 ± 3.82*	12.13 ± 3.61*	4.00 ± 0.96*
MDA-MB-231 Spheroid (3D)	37.72 ± 5.31 *	8.083 ± 2.44*	15.67 ± 1.71*

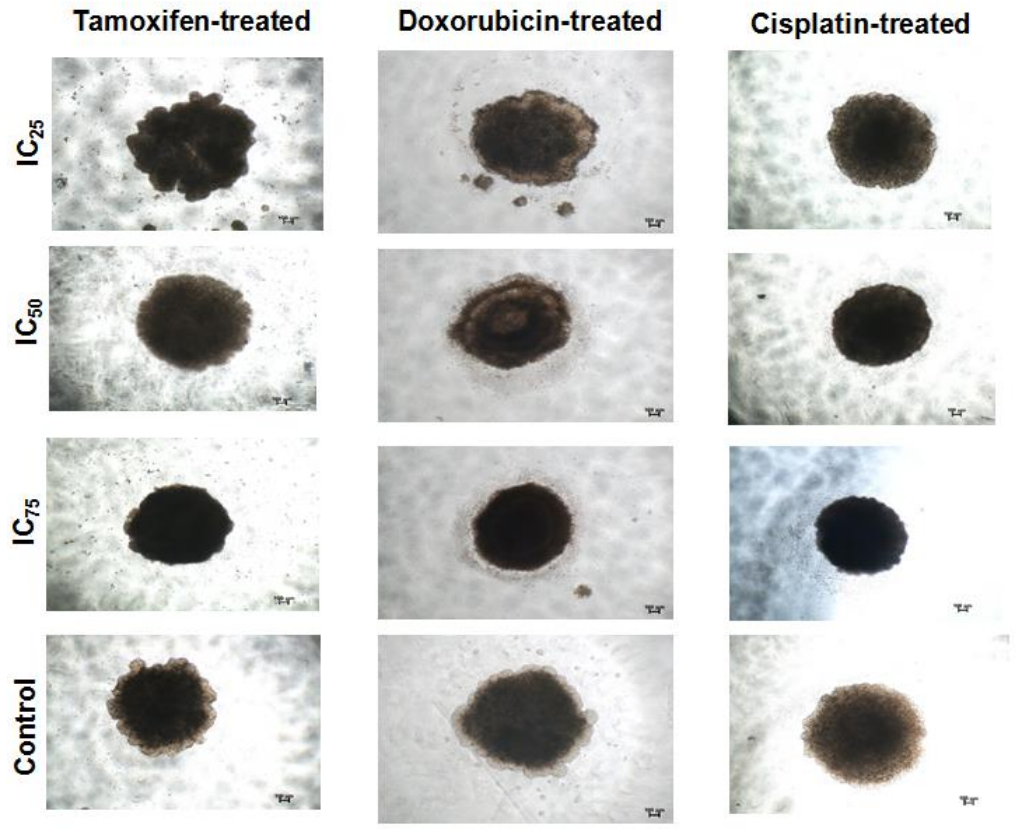


Figure 4.29. Spheroid integrity following treatment with chemotherapeutic drugs at different inhibitory concentrations of MCF-7 cells (3D condition). The morphology of the spheroids at 96 hours post-treatment were captured using bright field microscopy. All images were at 40X magnifications.

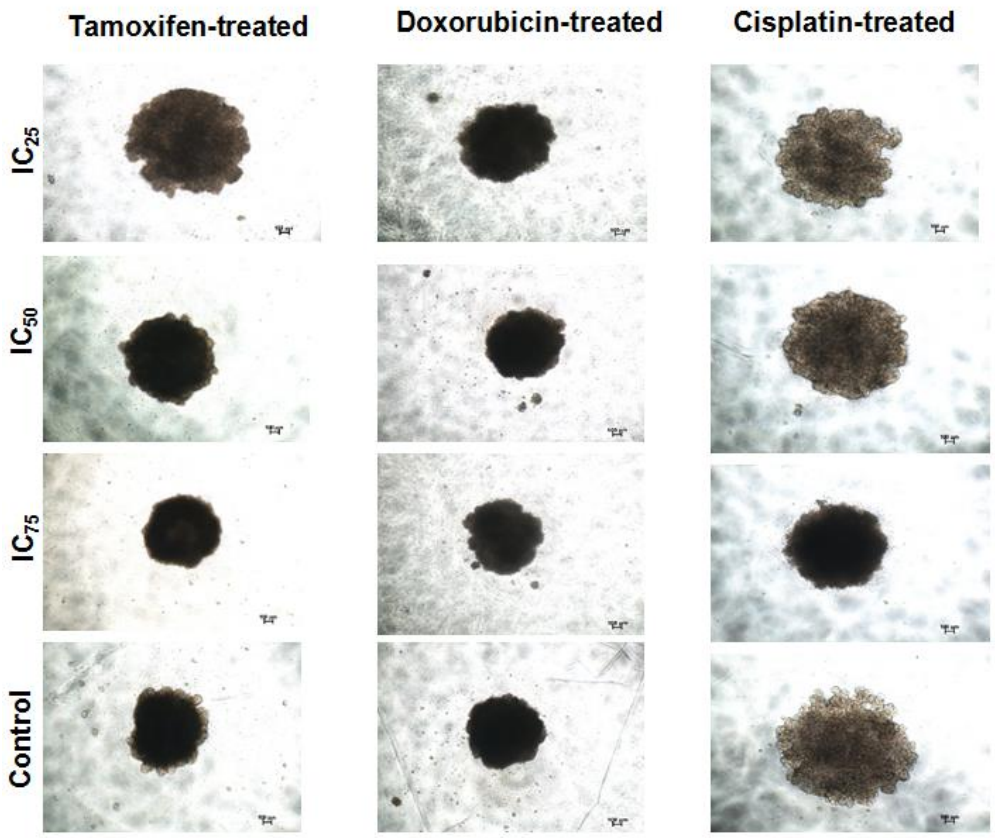


Figure 4.30. Spheroid integrity following treatment with chemotherapeutic drugs at different inhibitory concentrations of MDA-MB-231 cells (3D condition). The morphology of the spheroids at 96 hours post-treatment were captured using bright field microscopy. All images were at 40X magnifications.

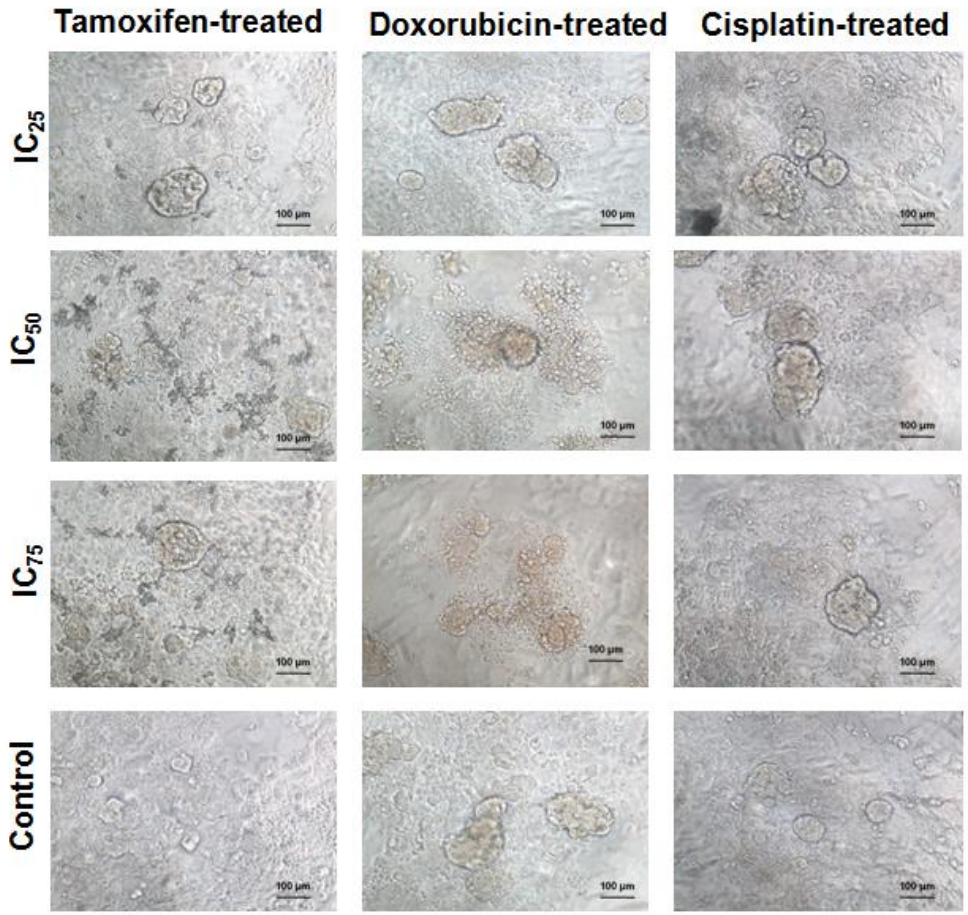


Figure 4.31. Morphological changes of MCF-7 cells (2D condition) following treatment with chemotherapeutic drugs at different inhibitory concentrations. The morphology of the cells at 96 hours post-treatment were captured using bright field microscopy. All images were at 100X magnifications.

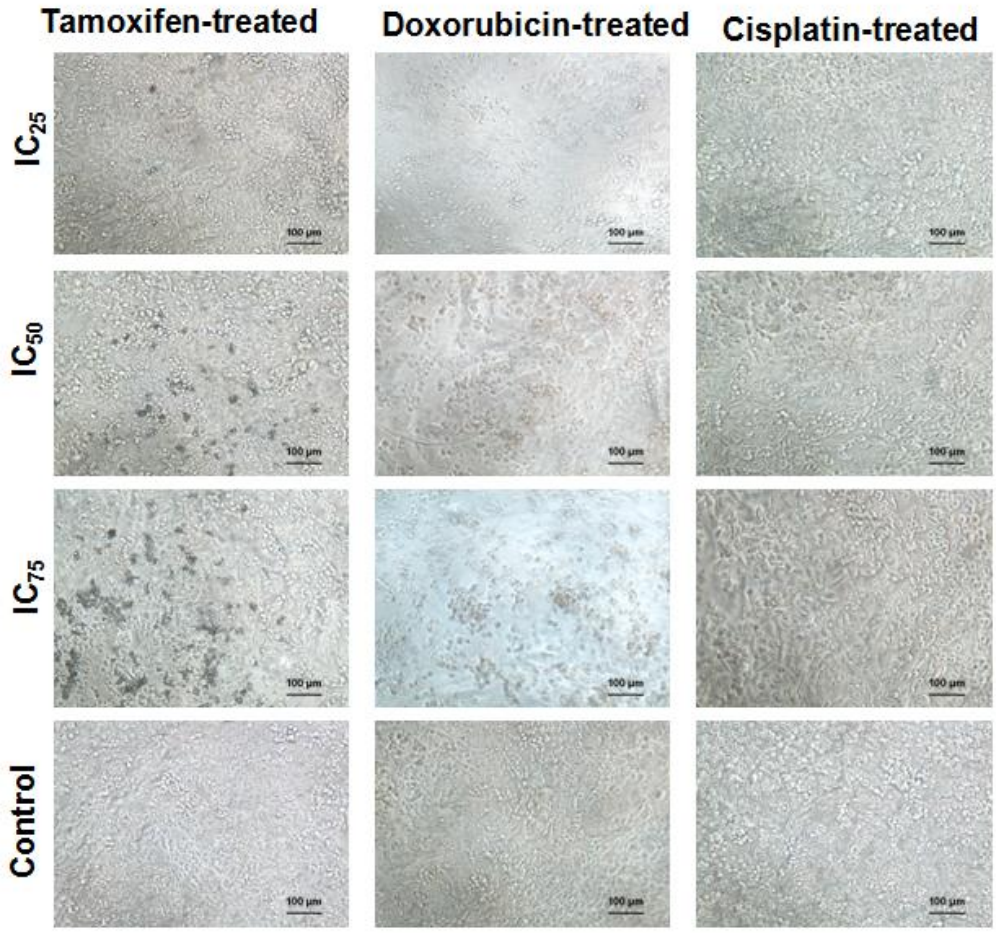


Figure 4.32. Morphological changes of MDA-MB-231 cells (2D condition) following treatment with chemotherapeutic drugs at different inhibitory concentrations. The morphology of the cells at 96 hours post-treatment were captured using bright field microscopy. All images were at 100X magnifications.

.

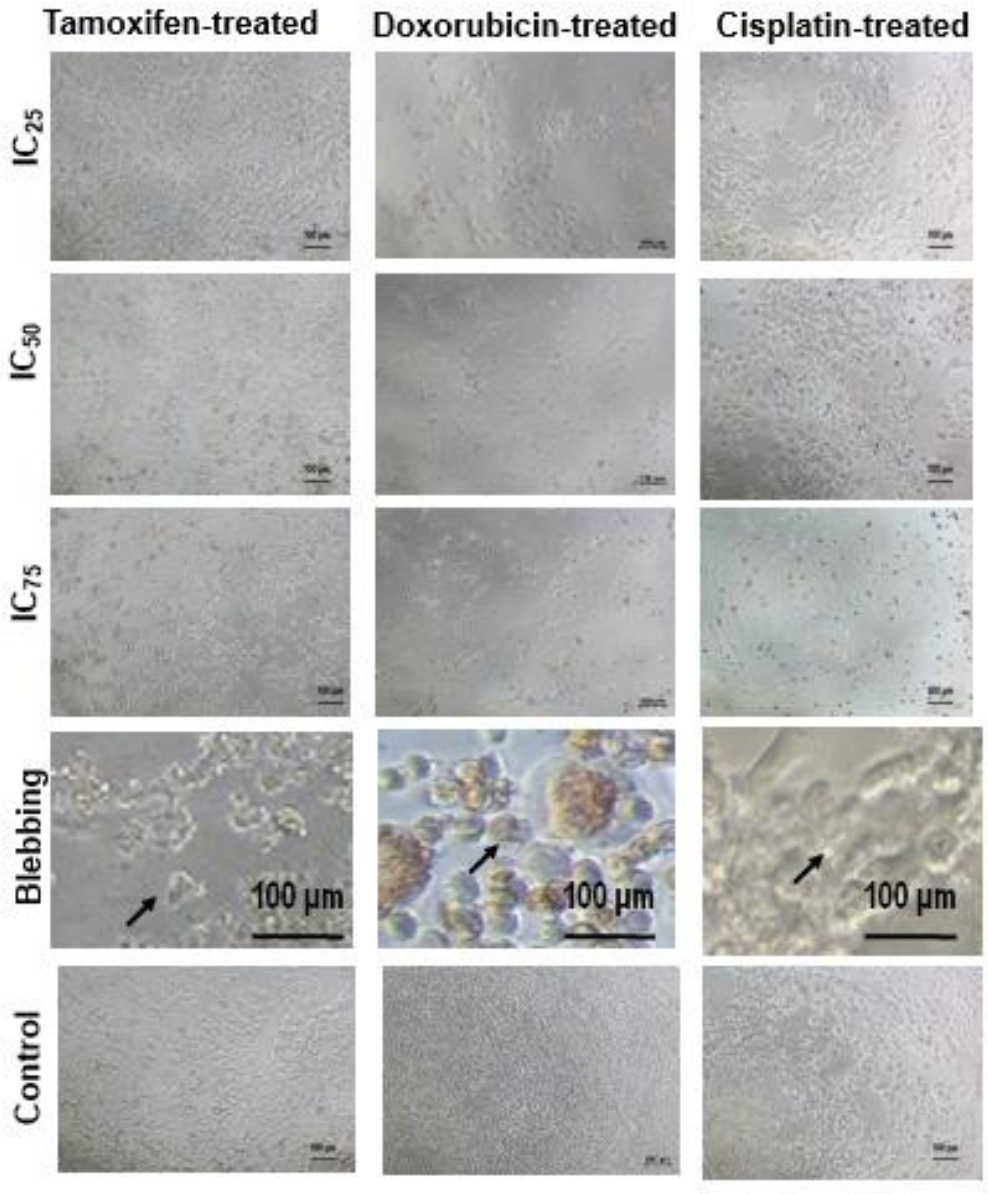


Figure 4.33. Morphological changes of MCF-7 parental monolayer cells following treatment with chemotherapeutic drugs at different inhibitory concentrations. MTT assays were performed following drug treatment protocol for monolayer cultures for the parental cells. The morphology of the cells at 96 hours post-treatment were captured using bright field microscopy. Arrows indicate membrane blebbing as evidence of apoptosis following drug treatment. Membrane blebbing images were at 400X magnifications, others images were at 100X magnifications.

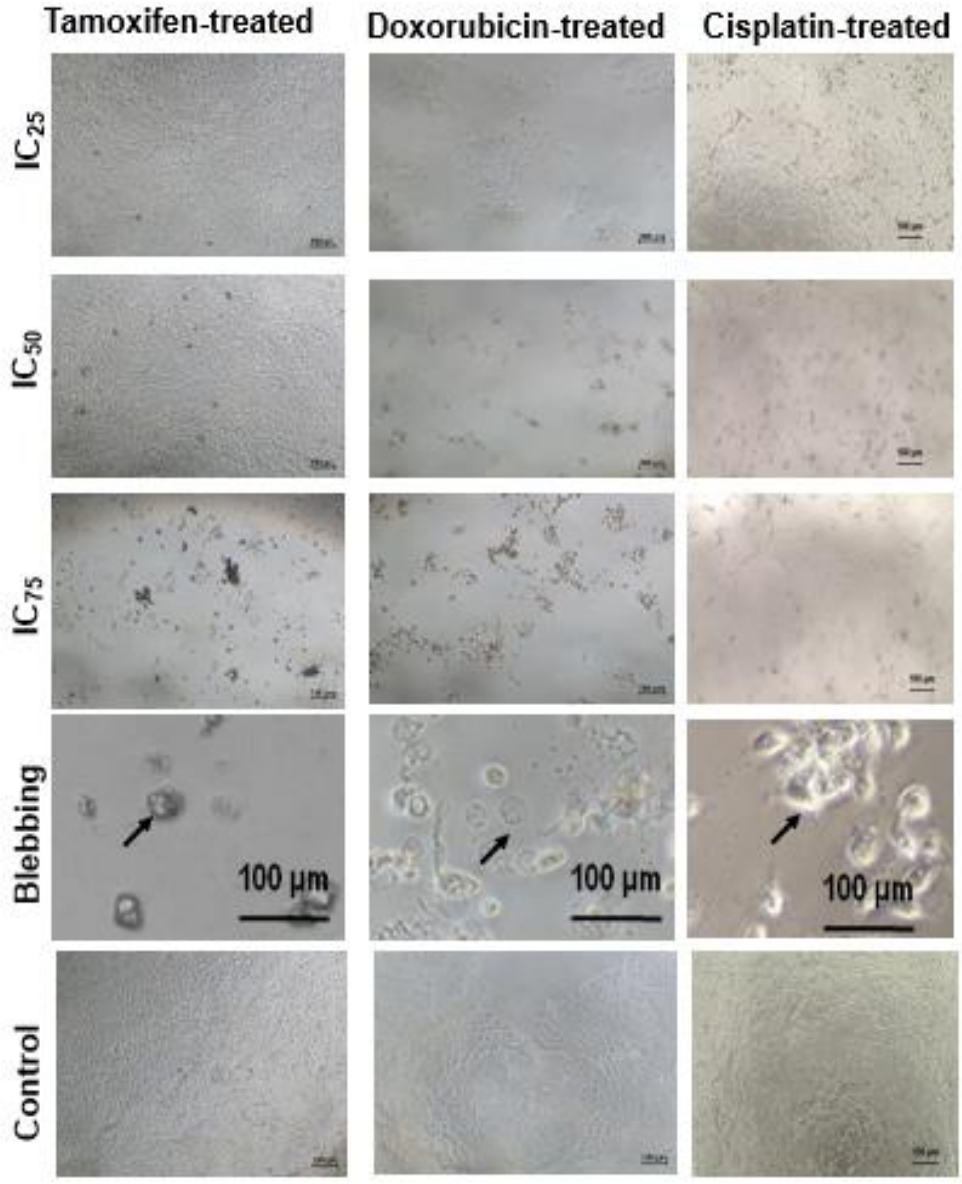


Figure 4.34. Morphological changes of MDA-MB-231 parental monolayer cells following treatment with chemotherapeutic drugs at different inhibitory concentrations. MTT assays were performed following drug treatment protocol for monolayer cultures for the parental cells. The morphology of the cells at 96 hours post-treatment were captured using bright field microscopy. Arrows indicate membrane blebbing as evidence of apoptosis following drug treatment. Membrane blebbing images were at 400X magnifications, others images were at 100X magnifications.

. 4.3.8 Higher Cell Proliferation Rate in Spheroid Cells

In order to determine the cell proliferation rate of the spheroid cells when compared to its parental cells, alamarBlue (AB) cell proliferation assay was performed. AB assay is a colorimetric growth indicator based on detection of metabolic activity that incorporates an oxidation-reduction in response to the chemical reduction of growth medium resulting from cell growth (Rampersad, 2012). The length of incubation time and the plating density for the assay were determined from the linear range of the graphs (Figure 4.35). The desired incubation time were 4 hours and the range of plating density was found to be from 200 to 500 cells/well for both cell lines.

Cell proliferation assay was observed to be higher in dissociated cells derived from spheroid compared with the parental cells (Figure 4.36A and Appendix G). Increased in cell size was observed in dissociated spheroid cells from day 3 and continued to increase in size to day 14 as shown in Figure 4.36B. Since the cell seeding density and culture conditions were the same, the varying cell proliferation rate observed could be due to higher amount of enriched CSCs in the spheroid cells as compared to the parental cells.

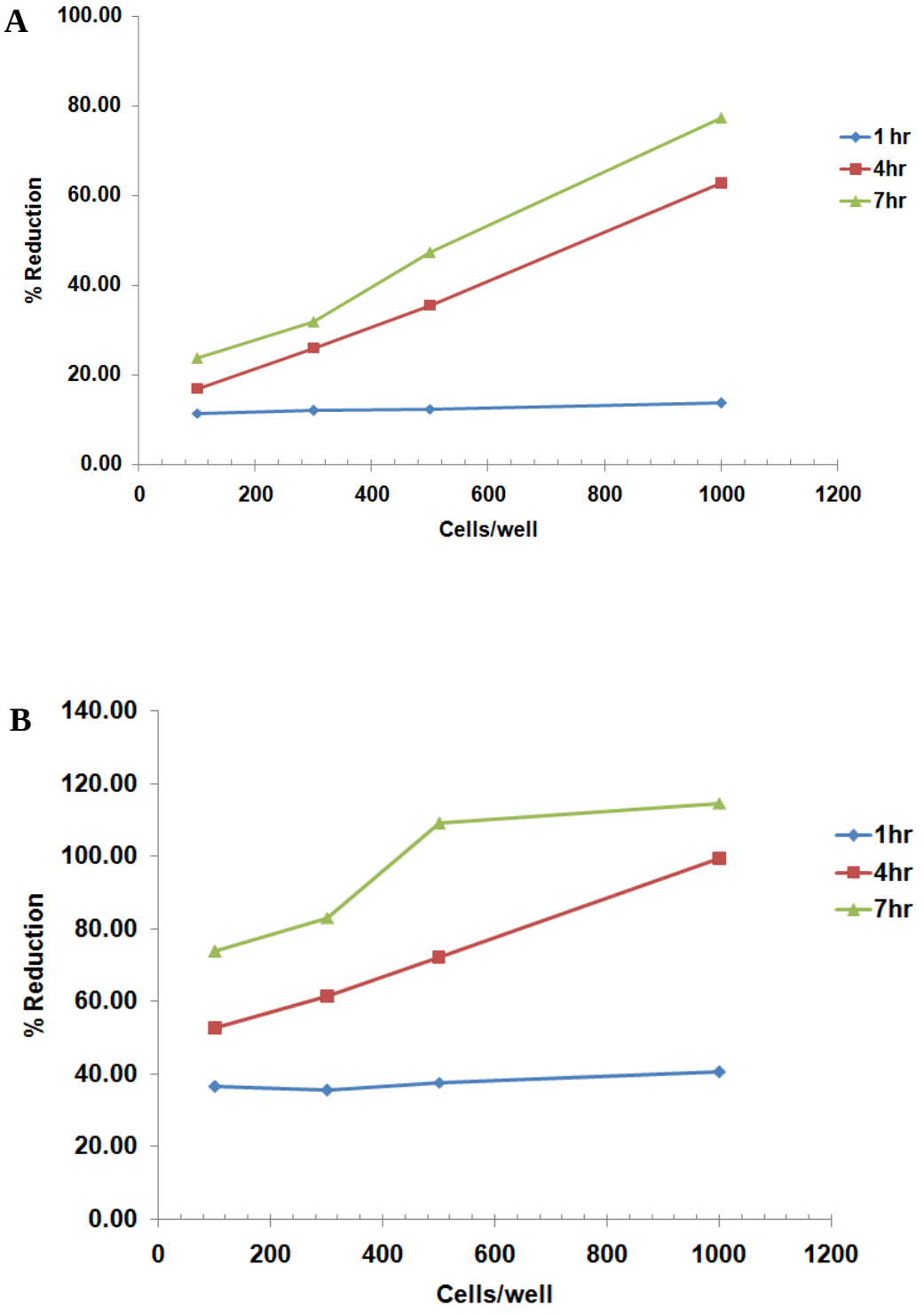


Figure 4.35. Percentage of alamarBlue (AB) reduction for MCF-7 and MDA-MB-231 monolayer cultures with different initial cells per well and incubation time. (A) MCF-7 cells. (B) MDA-MB-231 cells.

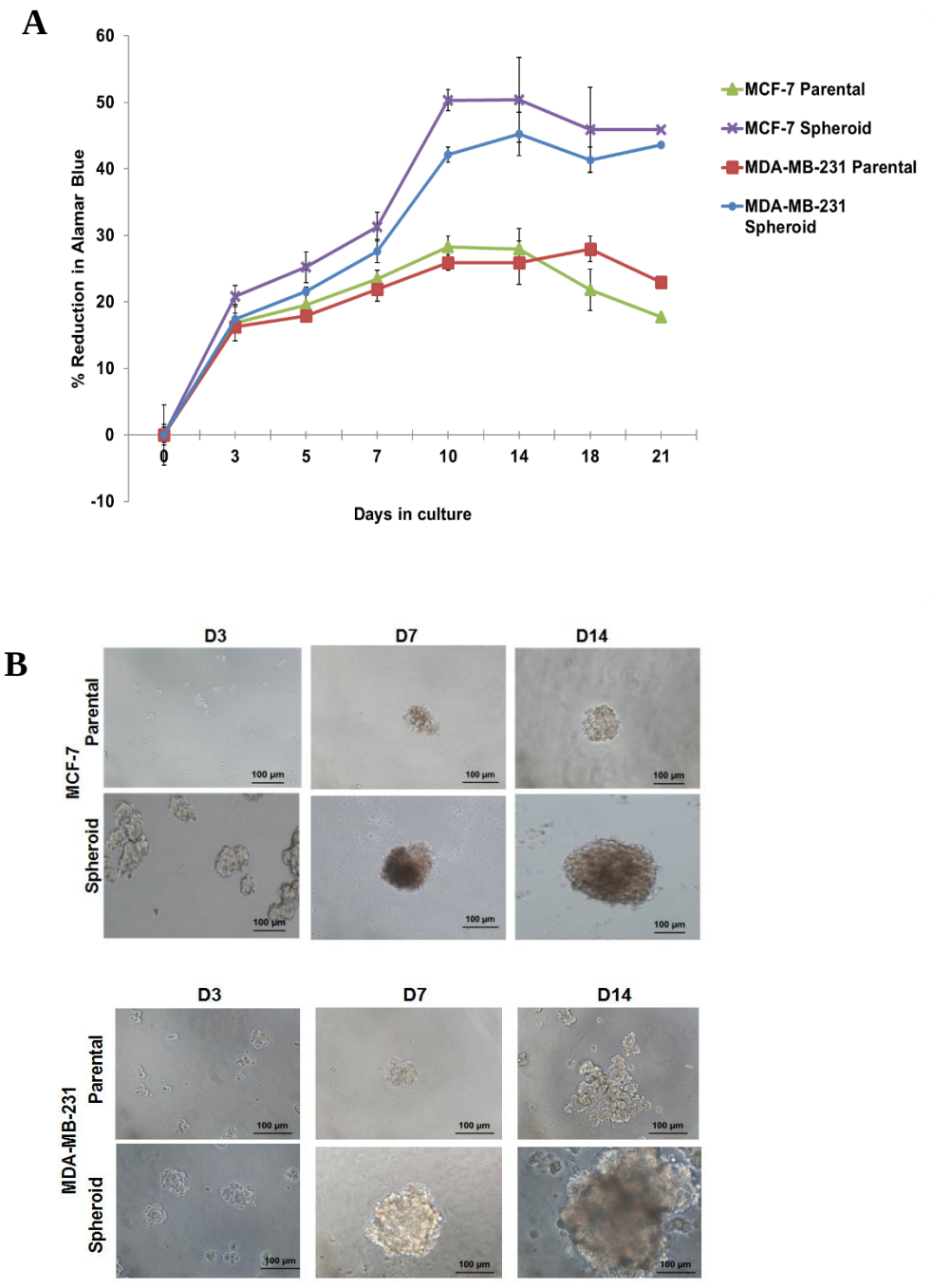


Figure 4.36. Comparison of cell growth of spheroid and parental cells of MCF-7 and MDA-MB-231 breast cancer cell lines. (A) Spheroids demonstrated a higher cell proliferation rate as compared to their parental cells based on the higher % of reduction of alamarBlue. (B) Morphological changes of the spheroid formation from dissociated cells derived from parental and spheroid cells. All images were at 100X magnifications.

4.3.9 Higher Wound Healing, Migration and Invasion Ability in Spheroid Cells

In order to examine the wound healing, cell migration and invasion of the spheroid cells relative to the parental cells, three assays were performed. In wound healing assay (scratch assay), spheroid cells on the edge of the newly created "gap" migrated faster to close the wound compared to the parental cells (Figure 4.37A and 4.38A). Statistically higher wound healing rate were observed in spheroid cells than in their parental groups (Figure 4.37B and 4.38B). Three independent experiments for each of the cell type is outlined in Appendix H. The observation that the spheroid cells migrated quicker than the parental counterparts was further confirmed by the cell migration assay. Compared with the parental cells, the migratory ability of spheroids was higher as there were more cells passing through the trans-well membrane as indicated in Figures 4.39 and 4.40 (migration assay). Meanwhile, the invasiveness of spheroids cells was also found to be higher as there were more cells invading the matrigel after 24 hours of culture as shown in Figures 4.39 and 4.40 (invasion assay). Three independent experiments were carried out as outlined in Appendix I. Collectively, these results illustrated that CSCs subpopulations enriched in spheroid cells could enhance the proliferation, migration and invasion rates of spheroid cells *in vitro*, which accounted for the observed differences when compared to their parental counterparts.

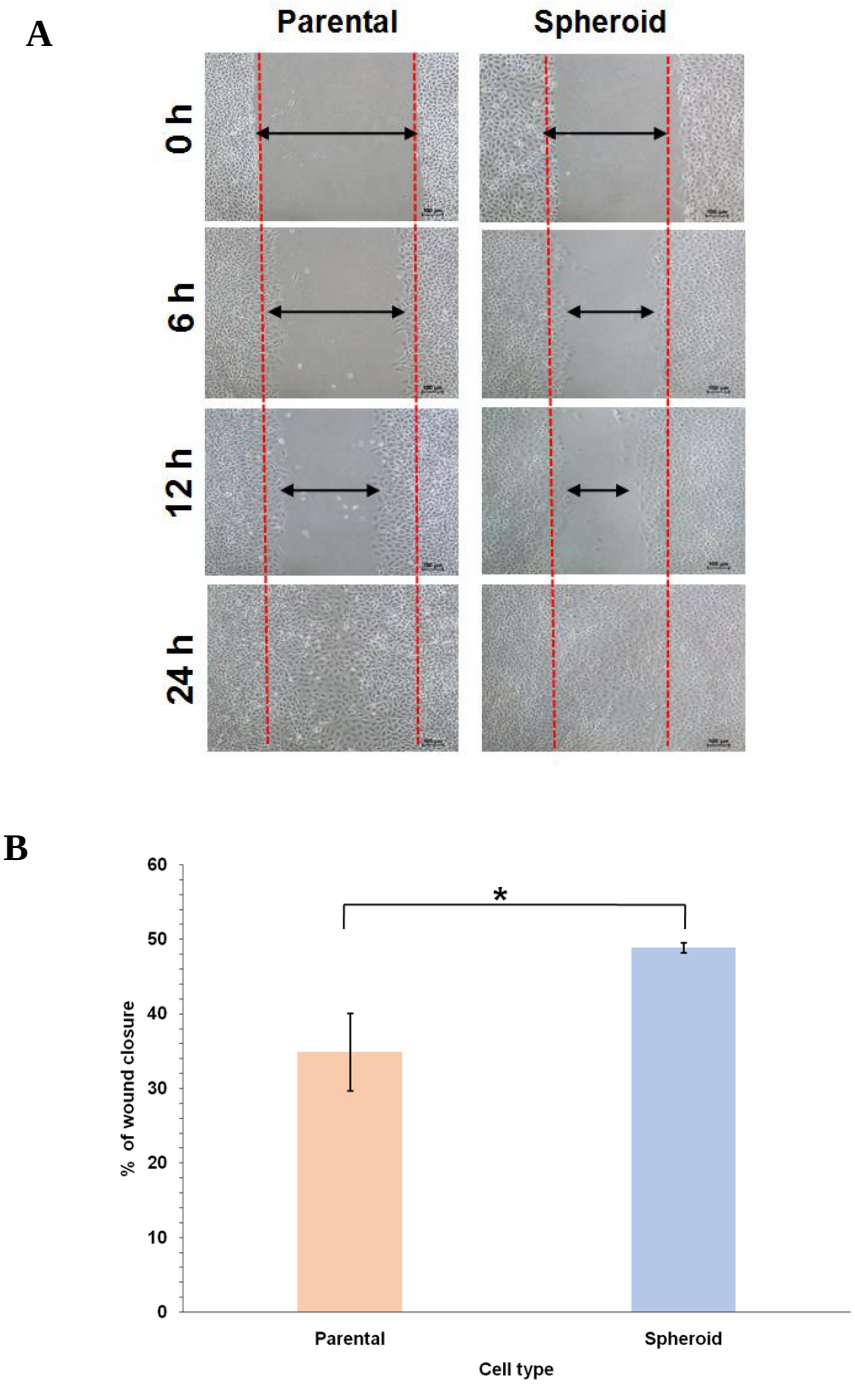


Figure 4.37. Wound healing assay of MCF-7 spheroid and parental cells. (A) Movement of cells into the wound from 0 to 24 hours. Red dotted lines indicated the boundary of the scratch assay and arrows indicated the recovery of the gap by the migrated cells. (B) Results were expressed as percentage of wound closures of the different cell types. Data were the means of three measurements, and the bars were the SD of the mean. * p<0.05. All images were at 100X magnifications.

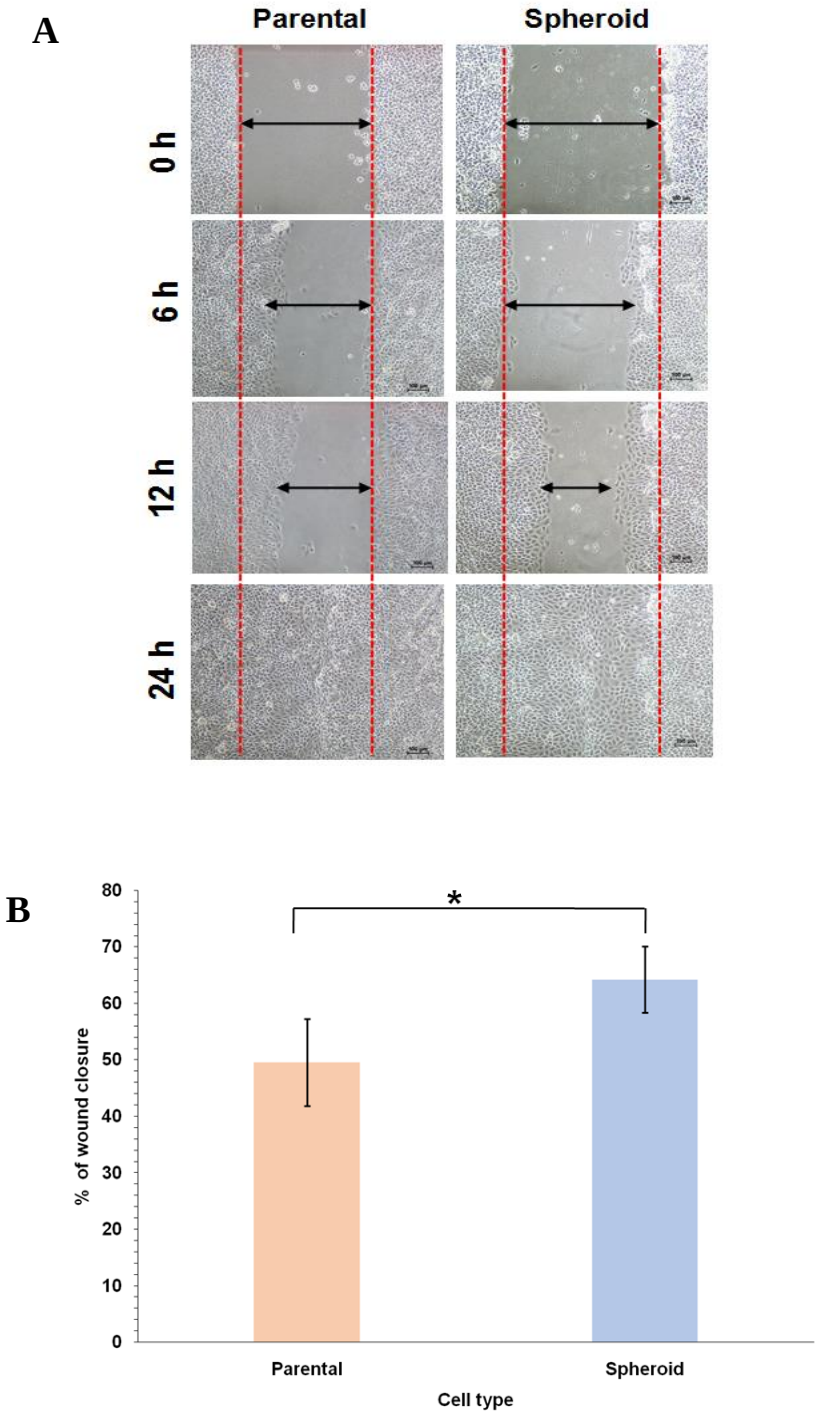


Figure 4.38. Wound healing assay of MDA-MB-231 spheroid and parental cells. (A) Movement of cells into the wound from 0 to 24 hours. Red dotted lines indicated the boundary of the scratch assay and arrows indicated the recovery of the gap by the migrated cells. (B) Results were expressed as percentage of wound closures of the different cell types. Data were the means of three measurements, and the bars were the SD of the mean. * p<0.05. All images were at 100X magnifications.

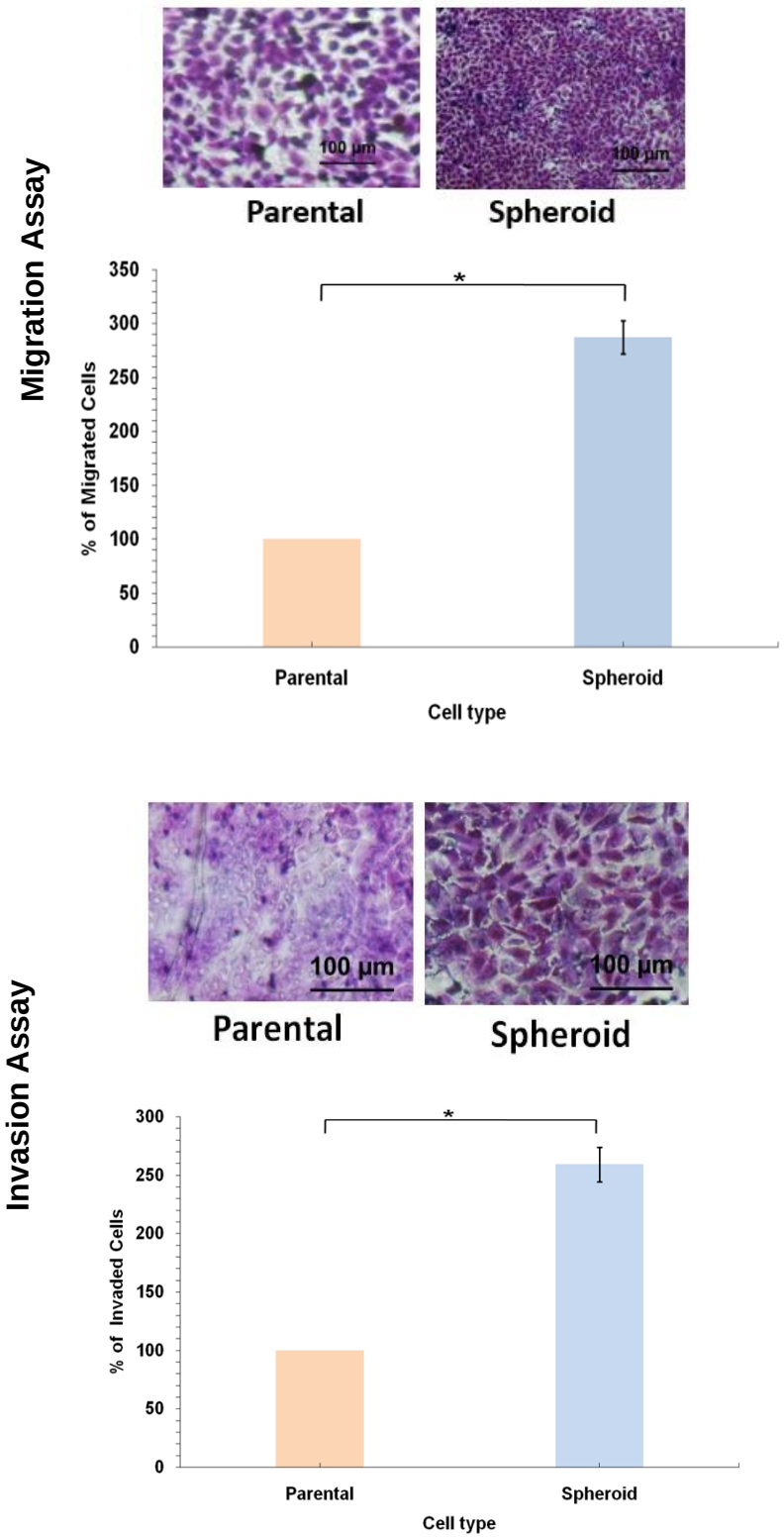


Figure 4.39. Migration and invasion assay of MCF-7 spheroid and parental cells. Shown here are the percentage of migrated/invaded cells of the spheroid and parental cells. Spheroid had a higher migration and invasion rate than their parental cells. Data were the means of three measurements, and the bars were the SD of the mean. *p<0.05. All images were at 100X magnifications.

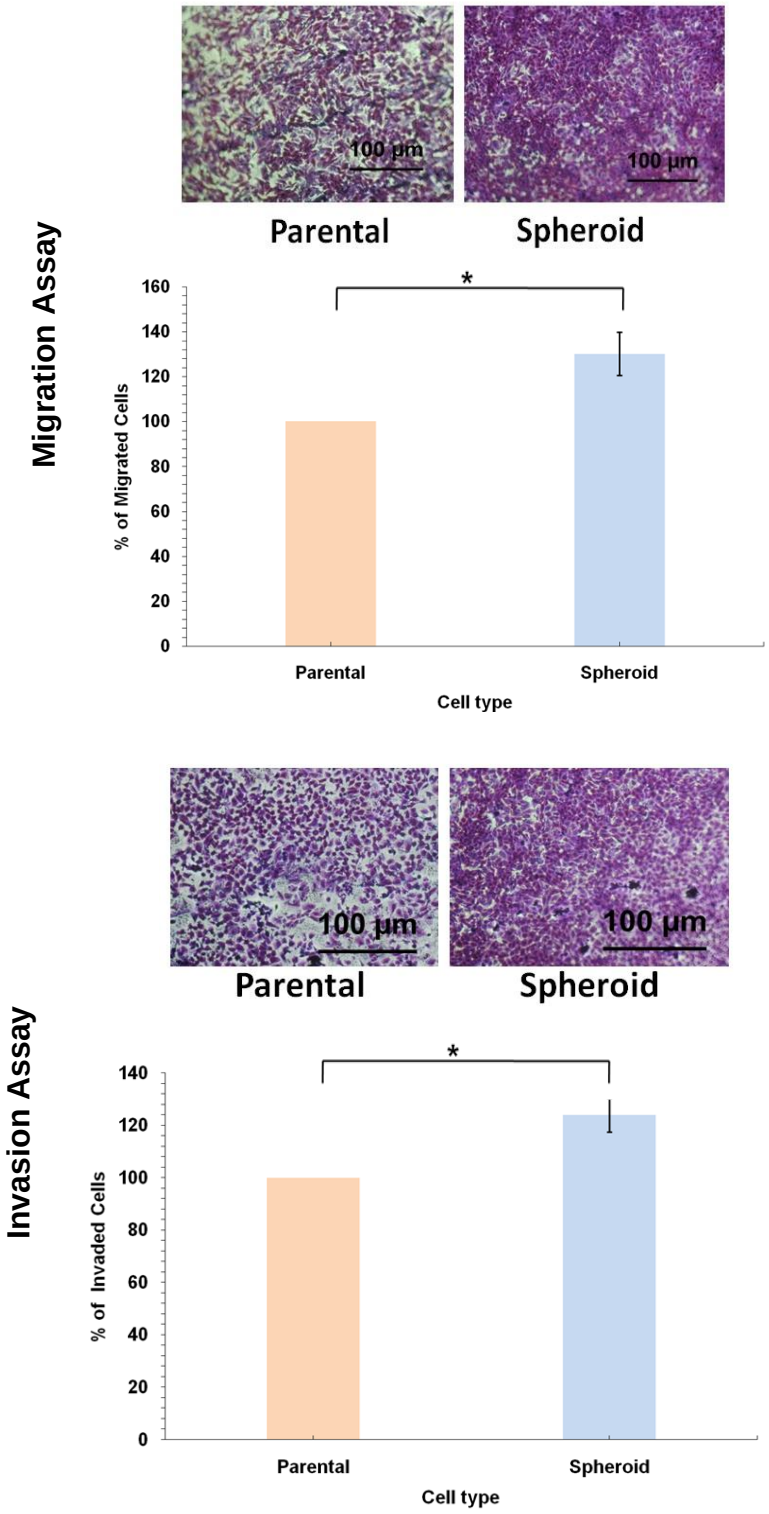


Figure 4.40. Migration and invasion assay of MDA-MB-231 spheroid and parental cells. Shown here are the percentage of migrated/invaded cells of the spheroid and parental cells. Spheroid had a higher migration and invasion rate than their parental cells. Data were the means of three measurements, and the bars were the SD of the mean. *p<0.05. All images were at 100X magnifications.

4.3.10 Cell Cycle Profile in Spheroid Cells

To determine the cell cycle stage progression in parental and spheroid cell types of MCF-7 and MDA-MB-231, cell cycle flow cytometric analysis was performed by staining the DNA. It was observed that the MCF-7 spheroid cells had an increased percentage of cells in Sub G0/G1, S and G2/M phases, and reduction in G0/G1 phase (Figure 4.41 and Appendix J). The percentage of cells in the Sub G0/G1 phase was significantly increased from 1.02 ± 0.04% (parental cells) to 4.04 ± 0.26% (spheroid cells); in the S phase, the number of cells increased from 1.10 ± 0.07% (parental cells) to 7.05 ± 0.37% (spheroid cells) and; in the G2+M phase, the increment was from 6.92 ± 0.33% (parental cells) to 13.68 ± 0.65% (spheroid cells). Meanwhile, the percentage of spheroid cells in the G0/G1 phase was lower from the parental cells (77.12 ± 0.44% compared to 90.60 ± 0.26%). For the MDA-MB-231 spheroid, a different cell cycle profile was obtained with over half of the cell population constituted the G0/G1 phase (72.45 ± 0.45%) compared to parental cells (48.90 ± 0.34%). There was also increment of populations of cells in the Sub G0/G1in the spheroid cells relative to the parental cells (2.40 ± 0.15% versus 0.52 ± 0.17%). Meanwhile, the percentage of cells in the S and G2+M phase was decreased in the spheroid cells compared to their parental counterparts from 29.22 ± 0.26% to 7.86 ± 0.24% and 20.08 ± 0.16% to 16.90 ± 0.38%, respectively (Figure 4.42 and Appendix J). Overall, these data suggested that a different subpopulation of cells existed in spheroid cells compared to its parental counterparts, accounted for the different cell cycle profile.

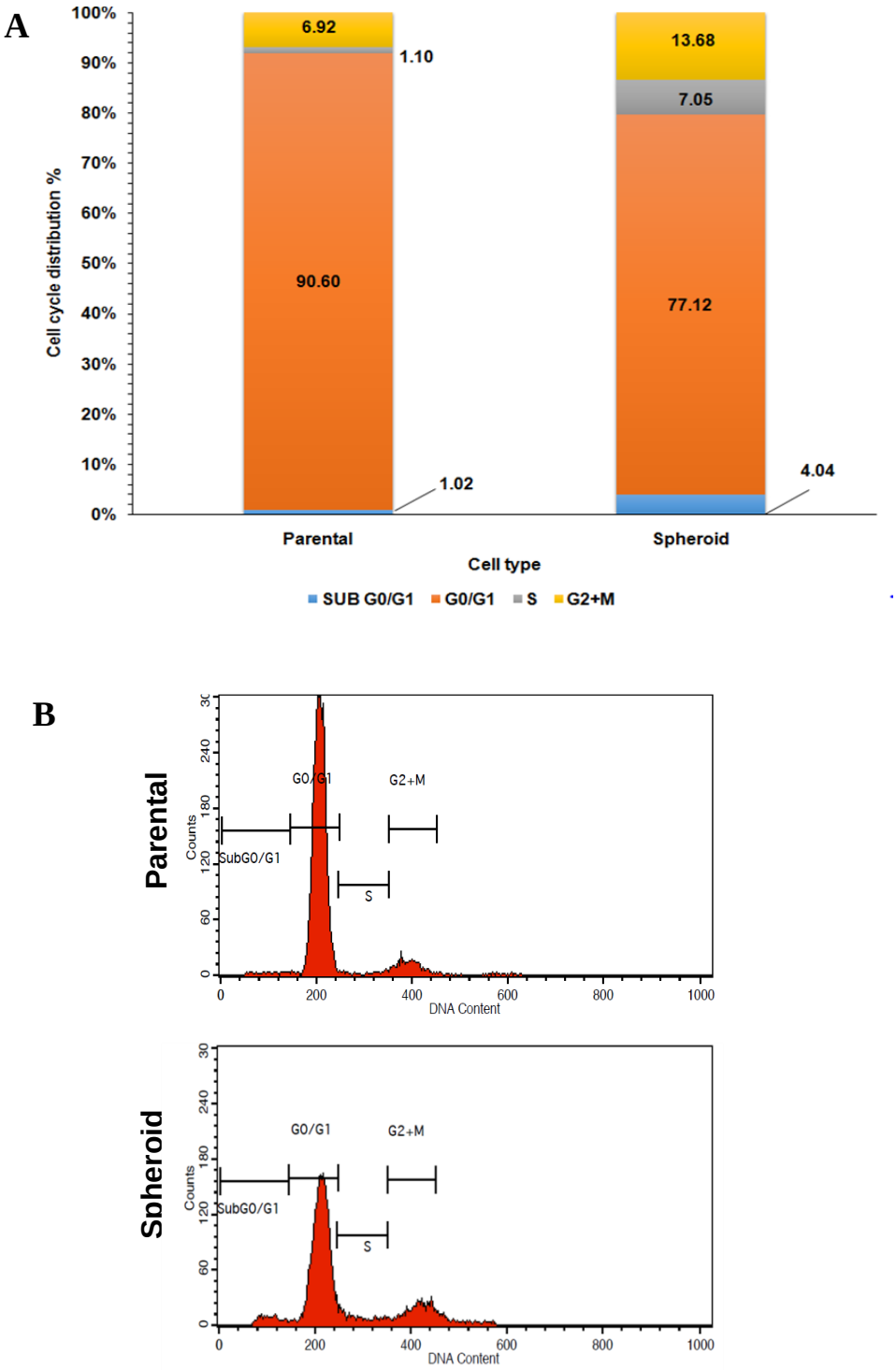


Figure 4.41. The proportion of cells at different stages in the cell cycle of spheroid and parental MCF-7 cell line. (A) Bar charts illustrating the percentage of cell cycle distribution (B) Representative flow cytometry histograms of the spheroid and parental MCF-7 cell cycle. The phases of the cell cycle: Sub G0/G1, G0/G1, S and G2/M are indicated on the histograms.

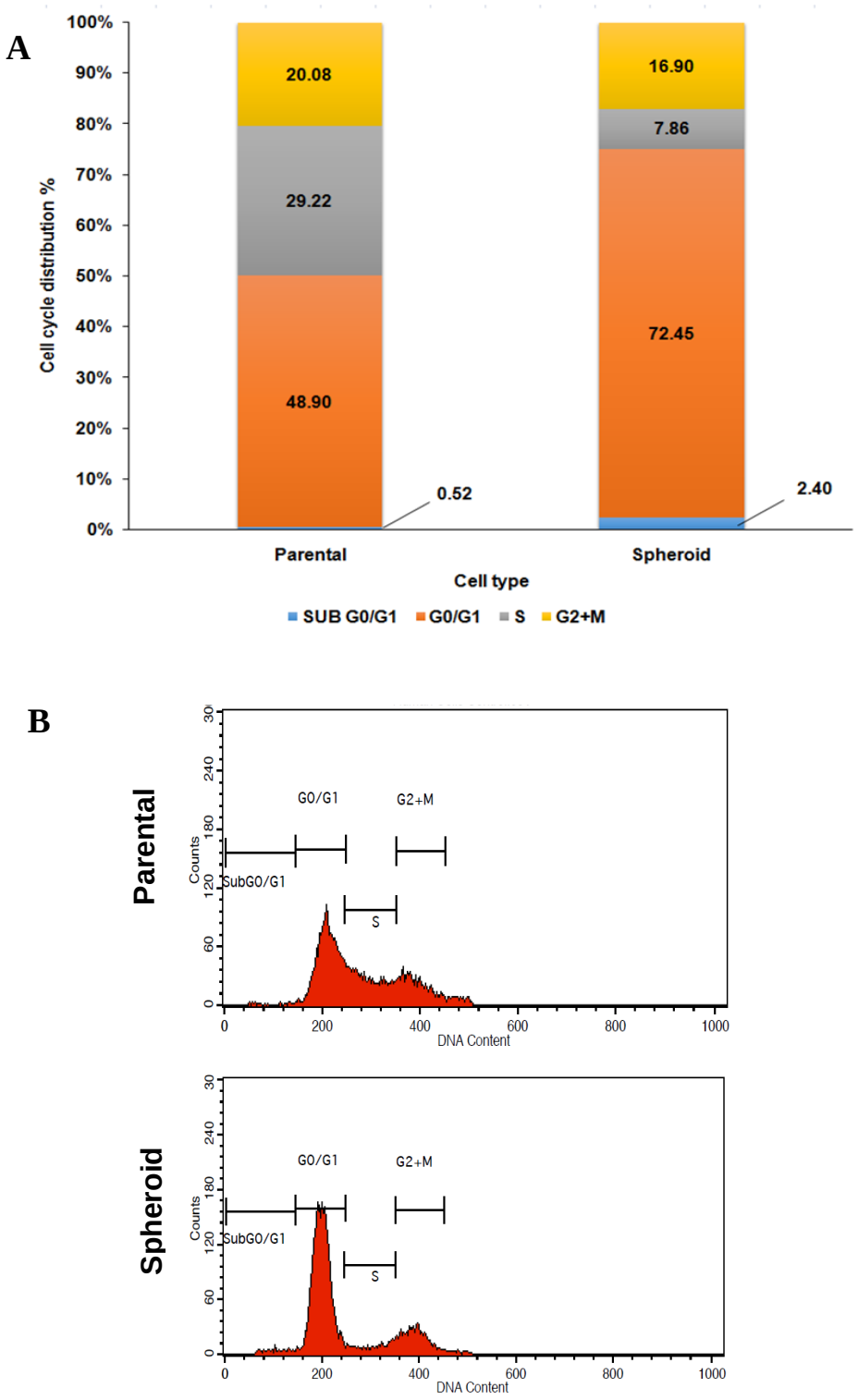


Figure 4.42. The proportion of cells at different stages in the cell cycle of spheroid and parental MDA-MB-231 cell line. (A) Bar charts illustrating the percentage of cell cycle distribution (B) Representative flow cytometry histograms of the spheroid and parental MDA-MB-231 cell cycle. The phases of the cell cycle: Sub G0/G1, G0/G1, S and G2/M are indicated on the histograms.

4.3.11 Spheroids Cells Exhibit DNA Hypomethylation Pattern

Global DNA methylation was assessed by an ELISA-like assay, using an antibody against 5-methylcytosine (5-mC), which detects cytosine methylation at both CpG sites and non-CpG sites. Prior to determine the methylation levels on the samples, standard curve with 5-mC was generated. The percentage of hydroxylmethylated DNA was proportional to the absorbance intensity measured, with the regression coefficient found to be 0.974 (Figure 4.43A). It was found that global methylation was significantly decreased in spheroid cells as compared to the parental cells, suggesting that enriched spheroid cells leads to DNA hypomethylation (Figure 4.43B and Appendix K).

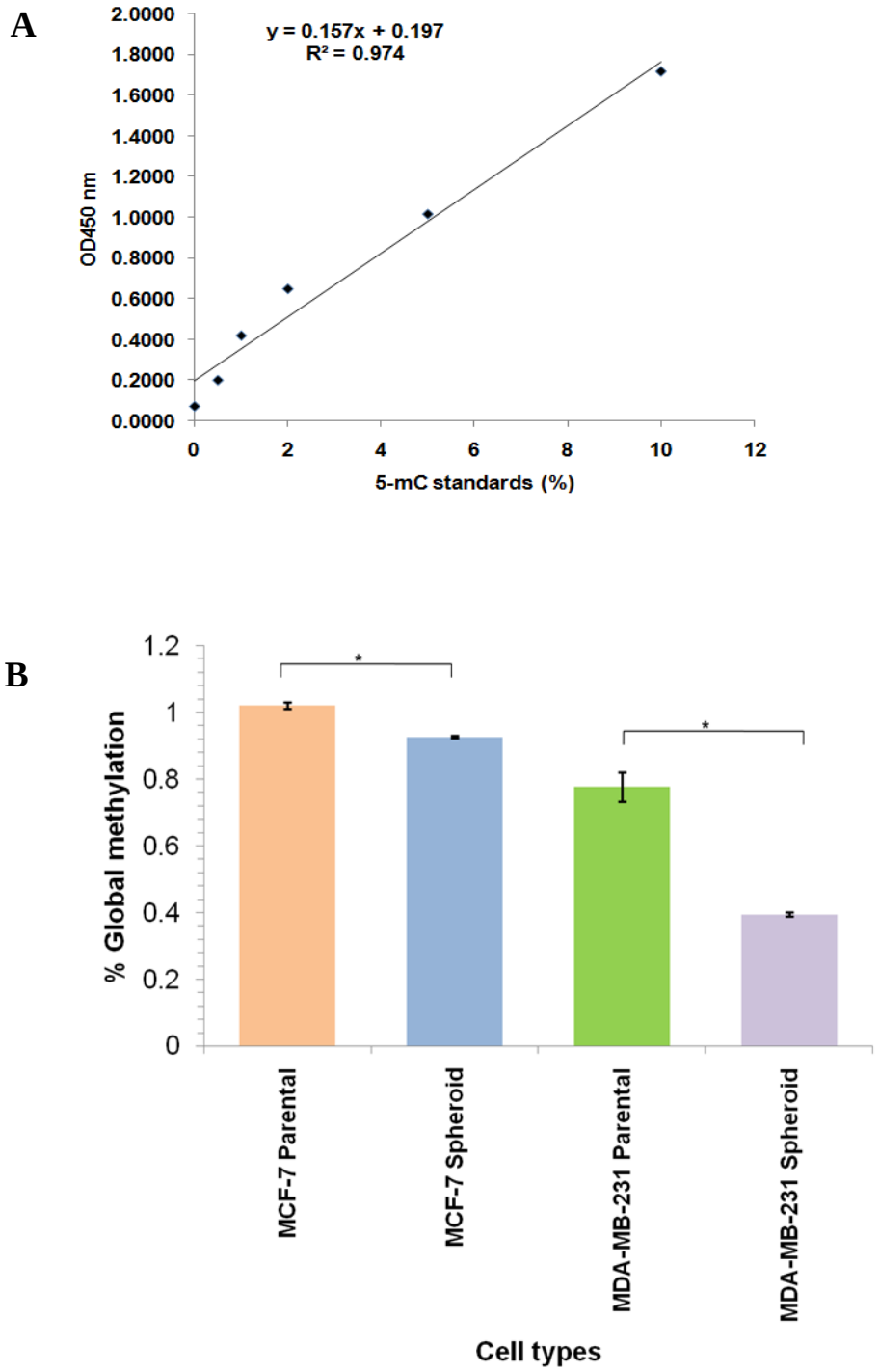


Figure 4.43. Quantification of 5-methylcytosine (5-mC) content of the DNA samples from the parental and spheroids cells. (A) Generation of the standard curve with known 5-mC controls. (B) Percentage of the global methylation (5-mC/Total DNA) using the MethylFlash Colorimteric DNA methylation kit. Data were the means of three measurements, and the bars were the SD of the mean. *p<0.05.

4.4 Discussion

4.4.1 Spheroid Culture Enrich CSCs Subpopulations

In this study, an agar overlay technique and centrifugation force was employed to generate the tumour spheroid from human breast adenocarcinoma oestrogen-dependent MCF-7 and oestrogen-independent MDA-MB-231 cells in serum-free culture condition. First, unlike other methods, our methods utilizing the concave agar surface and combining with centrifugal force allow the cells to seed on the agar surface and form spheroids that were homogenous in size. The cells aggregated on the agar matrix allowing the monitoring of the morphological changes in terms of their formation and progression. The centrifugal force and the concave surface of the agar matrices used in this study allowed the cells to form spheroids that were homogeneous in size. Other matrices that enable the growth of cells in spheroids have been also reported, they are namely matrigel, collagen, alginate, chitosan and polyHEMA (poly(2-hydroxyethyl methacrylate) (Langhans, 2018, Ray, Morford et al., 2018). The wide variety of these commercially available matrices is expected as different matrices are catered for different cell lines and also their endpoint applications. Among all the matrices, the use of agarose is preferred in our study because it is inexpensive, easy and does not involve specialised skills or laboratory equipment for its preparation (Guo, Chen et al., 2019). Nevertheless, all these matrices are meant to create a non-adherent culture condition to promote the 3D spheroid cell growth.

Secondly, our method was able to enrich the CSCs population using the serum-free media as described previously (Mehta, Novak et al., 2018). During the period of the sphere cultures in serum-free condition, this culture condition is thought to keep CSCs in their undifferentiated stage facilitating their enrichment. Although deprived of serum, the cells utilise their own extracellular matrices and organise in a complex cell-to-cell interaction to form a heterogeneous spheroid, mimicking the tumour tissues *in vivo* (Hirschhaeuser, Menne et al., 2010, Nagelkerke, Bussink et al., 2013). The tight and compact spheroid formed in this study suggested a different intrinsic mechanism compared to their monolayer parental cells. The decrease in cell number of the "mature" spheroid compared to the initial cell density demonstrated cell loss, which likely due to the unfavourable serum-free culture condition. The serum-free culture condition facilitates the enrichment of CSCs-like population within spheroids of both subtypes, with observable initial cell loss and reduction in size. This is no surprising as reduction in cell number could be due to the restricted oxygenation and nutrition diffusion in the 3D spheroids as compared to 2D cultures where the environment is rich in oxygen and nutrients (Gomes, Guillaume et al., 2016). Confocal imaging on 3D spheroids has also revealed the different layers of cells representing the typical zones of cell proliferation within a 3D spheroid (Murali, Chang et al., 2019). Nonetheless, serum-free medium was thought to keep the CSCs-like cells in their undifferentiated stage whereas the fast proliferating cells located in the outer region of the spheroid could have succumbed to death due to their intolerance to the serum-depleted environment (Galli, 2019). This is in agreement in other published studies in which they hypothesised that serum-

free medium could effectively enrich CSCs subpopulations as only CSCs were able to survive and proliferate in serum-deprived condition (Kruyt and Schuringa, 2010, Lin, Lee et al., 2012). The rest of the cells, which were speculated to be non-CSCs cell, may possibly undergo anoikis due to the unfavourable culture condition. However, some studies disagreed with this notion and indicated that not all cancer cells can enrich CSCs-like cells in spheroid culture and this method need to be used subjectively. Previously, Wu et al. in 2014 demonstrated that spheroid cells derived from Caco-2 colorectal cancer cell line did not consist of colorectal CSCs as they showed reduced specific colorectal CSCs surface markers, decreased chemoresistance ability and were not tumorigenic when injected into nude mice model. In another study, PC3, a prostate cancer cell line successfully formed spheres and demonstrated an increase of prostate CSCs surface markers and resistance to cisplatin. On the contrary, LNCaP, which was another prostate cancer cell line being investigated failed to survive and formed spheres in the non-adherent suspension culture (Fan, Liu et al., 2010). Therefore, spheroid enrichment technique for CSCs-like cells does not apply to all cancer cell lines, but in a cell line-dependent manner. Our initial assessment somehow revealed that the tumour spheroid model established using our method could be enriched with a subpopulation of CSCs based on the cell loss and reduction in size.

The ultrastructure of the spheroids performed using SEM analysis revealed that MCF-7 spheroid was made of dense cells with strong cell-to-cell adhesion within the spheroids that preserve its 3D shape. On the other hand, MDA-MB-231 spheroid appeared to be less compact with elongated cells and

presence of more pores as compared to the MCF-7 spheroids. MCF-7 is molecularly classified as Luminal A (ER+/PR+/HER2-) and corresponds to a relatively less aggressive breast cancer subtype. This could possibly result in the mass spheroid morphology. On the other hand, MDA-MB-231 which is a triple-negative basal-like (ER-/PR-/HER2-) and corresponds to molecularly mesenchymal-like and more aggressive phenotype formed loose spheroid. Hence, the nature of its subtype accounted for the differences in the spheroid morphology was also highlighted in other published works (Bianchini, Balko et al., 2016, Denkert, Liedtke et al., 2017). This observation was also noted in breast cancer cell lines compared with the cells from primary breast tumours they tested; differences in the spheroid morphology and efficiency was observed (Tamura, Sugiura et al., 2017). Different spheroid morphology was also observed in other types of cancer, in colorectal cancer cell lines for example, where it was found that the cell lines from the primary tumours formed 'mass' type spheroid, whereas cell lines derived from metastases exhibited a 'grape-like' morphology (Luca, Mersch et al., 2013). Previous work also highlighted the differences of the spheroid formation ability observed in the 11 breast cancer cell lines is associated with their expression of E-cadherin, a major protein involved in the epithelial to mesenchymal transition (EMT) (Liu, Gu et al., 2016). Overall, the morphological appearance of 3D spheroids is primarily cell line- and also cancer subtype-dependent.

These complex multicellular three-dimensional spheroids are not just a tool for understanding disease progression, or useful in drug screening, but may have important implications for CSCs disease modelling. Altogether,

spheroid cultures provide an easier approach to enrich and study CSCs-like cells in different cancers by means of reproducibility and validity of findings. It is also agreed that extensive characterisation of CSCs properties in spheroid derived from any cancer cell lines or cancer tissues must be carried out to ensure that the generated spheroids are actually enriched with CSCs. The expression of stem cells related markers, self-renewal capacity, resistant to drugs, metastatic capacity and miRNA profiling of these spheroid cells were still unclear.

4.4.2 Standard Phenotypic Characterisation of Spheroid-Enriched CSCs Models

The ability of the cells to grow as 3D spherical structures was further tested for the presence of self-renewing subpopulations of cells, in the secondary spheroids formation assays. Single cell and also 200 dissociated cells per well dilutions assays from the primary 3D spheroids of both cell lines demonstrated the formation of clonal, non-adherent secondary spheroids. Again, the ability of the spheroids to be serially passaged provides definitive evidence for the presence of self-renewable and extended proliferative cells. The preservation for self-renewal and extended proliferation properties in spheroid cells were consistent with other studies (Cao, Zhou et al., 2011, Liu, Wang et al., 2013). Ability to form secondary spheroids was also indicated in previous studies across different types of cancers enriched for CSCs (Kim, Pearson et al., 2016). Secondary spheroids from both cell lines were able to passaged up to three generations although differences could be observed (Figure 4.10A and Appendix B & C) corresponding to the breast cancer

subtypes. MDA-MB-231 secondary spheroids averagely enlarged into approximately 380 µm while MCF-7 secondary spheroids were estimated to be 303 µm in size after cultured for 14 days. Sphere-forming efficiency (SFE) analysis demonstrated that dissociated single cells from MDA-MB-231 spheroids had highest SFE rate as compared to MCF-7 spheroids. MDA-MB-231 cells characterised by aggressive features had a higher SFE and phenotypically larger could be due to the higher proportion of self-renewing cells than the MCF-7 cells (Shi Y et al., 2017).

We also further investigated the stemness characteristics using classical surface markers reported for breast CSCs identification (Fillmore and Kuperwasser, 2007, Grimshaw, Cooper et al., 2008). According to the CSCs criteria, cultured breast CSCs should exhibit a distinctive cell surface phenotype CD44+/CD49f+/CD24-/low. CD44 is a cell surface molecule that has roles in both adhesion and signalling and it has been associated with stem cell-like characteristics whereas CD24 is an epithelial molecule and associated with differentiated epithelial morphology (Martin, Harrison et al., 2003, Wei, Hu et al., 2012). It was also shown recently that cells with CD44+/CD24- phenotype in human mammary epithelial cells were associated with epithelial-mesenchymal transition (EMT). EMT process has been found to promote tumour invasion and metastasis in cancer cells through loss of contact inhibition, suggesting the roles of EMT in CSCs. (Morel, Lievre et al., 2008, Scheel and Weinberg, 2012). On the other hand, CD49f is an integrin receptor, which has been found to maintain pluripotency and stemness in embryonic

stem cells by regulating a variety of cellular activities (Yamamoto, Masters et al., 2012). Co-expression of CD44+ and CD49f+ has also been extensively study in prostate cancer that might represent a subpopulation of prostate CSCs (Guo, Liu et al., 2012). Therefore, immunofluorescences staining on the breast CSCs markers were used to evaluate the enrichment of CSCs in the spheroid cells relative to their parental cells. We found that CD44 and CD49f expressions was enriched in both spheroid cells compared with their parental cells as shown in the immunofluorescence staining, concurrent with other CSCs findings (Geng, Guo et al., 2013, Phillips, McBride et al., 2006, Ricardo, Vieira et al., 2011). The immunofluorescence staining expression pattern of the pluripotent transcription markers (Nanog and Sox2) on spheroid enriched cells of both subtypes was expressed as compared with the parental cells, indicating presence of pluripotency-like population of cells within the spheroids.

It was recently showed that increased of aldehyde dehydrogenase activity (ALDH) in cancer have stem/progenitor cell properties (Li, Su et al., 2010). ALDH1 is an enzyme associated with the production of retinoic acid and has been linked to stem cells populations of multiple myeloma (Douville, Beaulieu et al., 2008). ALDH1 was also a marker to differentiate between normal and malignant human breast stem cells (Ginestier, Hur et al., 2007). It has been reported that ALDH-positive cells demonstrated stem cell properties *in vitro* and were able to form tumours in NOD/SCID xenografts (Liu, Wang et al., 2016). On top of that, gene expression profiling identified a cluster of

related stemness genes in ALDH-positive cells (Blume, Rempel et al., 2018). ALDH1 activity based on fluorescent staining was found expressed in both spheroids, suggesting that ALDH1+ spheroid cells could be enriched with breast CSCs. To quantitate on the expressions of the two main breast CSCs surface markers (CD44 and CD24) as well as the ALDH intracellular enzyme activities in spheroid cells relative to their parental cells, flow cytometry was performed subsequently. The expression of CD44, CD24 and ALDH by flow cytometry indicated both spheroid cells had higher proportions of cells expressing those markers compared to their parental cells. Again, this shows that CSCs were gained upon serum-free enrichment culture on the two breast cancer cell lines. However, as we compared the CD44+/CD24- and ALDH+ expression between the spheroids subtypes, it was observed that MDA-MB-231 spheroids displayed a higher fraction of cells displaying those markers. This is in line with our earlier observation showing that MDA-MB-231 cells being a basal subtype of breast cancer are predominantly CD44+, accounted for their higher expressions of this surface marker. MDA-MB-231 spheroids also exhibited a higher ALDH expression, about 2-fold more than MCF-7 spheroid. Higher expression of ALDH in breast tumour cells has been linked with higher pool of CSCs populations, therefore it can be concluded that MDA-MB-231 spheroids were enriched with more CSCs population relative to MCF-7 spheroids. This assumption could also be explained by the fact that MDA-MB-231 characterised by their basal and mesenchymal phenotypes in which it has been shown to mask differentiation as compared to the increased expression of differentiation-associated genes found in luminal cell type (Jing, Liaw et al., 2016). Overall, increased expressions of CSC-related markers in spheroid

enriched CSCs suggested regulation of EMT-associated cancer progressions which is lacking in parental cells.

4.4.3 Other Phenotypic Characterisation of Spheroid-Enriched CSCs Models

In this study, both CSCs spheroids subtypes showed resistance to the three tested chemotherapeutic drugs in view of their higher drug inhibitory concentration relative to the parental cells (Table 4.3). Previous studies have speculated that the higher degree of drug resistance could be due to the increased proportions of CSCs in the spheroids (Phi, Sari et al., 2018, Prieto-Vila, Takahashi et al., 2017). Our chemoresistant spheroids enriched with detoxifying enzyme ALDH1 and CD44 glycoproteins attributed to the higher chemoresistance. Previous findings have confirmed that high expression of ALDH1 were associated with CSCs, and this marker shows promising potential to be used for the identification of CSCs (Mariel E. Toledo-Guzmán, Miguel Ibañez Hernández et al., 2019). It was also demonstrated that spheroid derived from osteosarcoma cell line MG63, positive for ALDH1 marker had strong chemoresistant capacities against both doxorubicin and cisplatin (Honoki, Fujii et al., 2010).

Meanwhile, CD44 expressions enriched in our spheroid may also account for the higher cell survivability as this marker involved in cell-cell and cell-matrix interactions, resulted in higher cell proliferation and migration (Liu,

Xiao et al., 2013). These two lines of evidence suggested that our chemoresistance spheroid expressing ALDH1 and CD44 had increased survival ability and tumourigenecity. Alternatively, it could also be explained by the drug barrier created by the 3D structure of the spheroids that conferred a higher degree of resistance to chemotherapeutic drugs as compared to the monolayer structures. Therefore, we tested out the drug resistance assays in the dissociated spheroid cells replated in monolayer condition to eliminate the possibility that the higher drug resistance in spheroids could be attributed by the physical barrier created by the many layers of cells in the spheroids (Nunes, Barros et al., 2018). Interestingly, at low drug treatment, IC25, smaller spheroids colonies were seen to be growing away from the main spheroid, indicating the metastasis ability of the spheroid cells. This has been observed in previous study where prolonged drug treatment enhances invasiveness of cancer cells and enrichment of CSCs (Branch, Masson et al., 2000, Calcagno, Salcido et al., 2010). We hypothesised that the reduction in spheroid volume is likely to be many factors. It could be attributed by the intrinsic mechanism of the heterogeneous populations of the spheroid cells and the microenvironment in which the spheroid cells reside in. Spheroids are arranged in a way that multilayer of cells interact with the extracellular matrix, resembling the arrangement of tumour *in vivo* (Carver, Ming et al., 2014, Ekert, Johnson et al., 2013). Therefore, being closely packed, spheroids develop a central necrosis and regions of hypoxia due to limited nutrient diffusion (Saggar, Yu et al., 2013, Tredan, Galmarini et al., 2007). The cytotoxic barrier and the cellular microenvironment have been identified as the contributing factors of drug resistance (Ivascu and Kubbies, 2007). Nevertheless, further investigations are

needed to confirm the data obtained from the microscopic images when spheroid cells are subjected for drug treatment. Moreover, these findings suggested that the increased proportions of CSCs could be the most probable contributor of the higher drug resistance in spheroids. Doxorubicin, one of the tested chemotherapeutic drugs, was more resistant in MCF-7 spheroids (18-fold) than MDA-MB-231 spheroids (8-fold) relative to their respective parental cells. MDA-MB-231 which formed a loosely and less compact spheroid than those of MCF-7, could be the reason that contributed to a higher doxorubicin drug penetration which resulted in a lower drug inhibitory concentration (IC_{50}). Likewise, a higher drug inhibitory concentration was also observed in MCF-7 spheroids than MDA-MB-231, in which the cytotoxicity of a ginger compound 6-shogaol in breast cancer cells and their respective spheroids was also reported (Ray, Vasudevan et al., 2015). In addition, enriched CSCs spheroid cells were less sensitive to the action of chemotherapeutic drugs could be due to the activation of human ATP-binding cassette (ABC) transporters. High expression of ABC transporters was associated with rapid removal of drugs from these cells (Herheliuk, T., et al., 2019). It has also been shown that higher ALDH enzymatic activity in enriched CSCs spheroid cells inactivates chemotherapeutic drugs, suggested the abundance of ALDH-positive cells was the cellular mediator for drug resistance (Liu, J. et al., 2013). Taken together, these results support a role for these spheroid cells enriched in CSCs in breast cancer chemoresistance, in which may explain why tumour tend to relapse after chemotherapeutic treatment.

Like in other study, a comprehensive characterisation to evaluate the phenotypic characteristics of the spheroid derived cells of both breast cancer cell lines for CSCs properties were performed. Accumulating evidences suggested that most of the cancer progression is associated with CSCs acquiring the epithelial to mesenchymal (EMT) phenotype, which is responsible for increased cell proliferation, motility and invasion (Shibue and Weinberg, 2017). Therefore, cells enriched for CSCs displayed greater proliferative, migratory and invasive potential when compared with non-CSCs cells. The rate of cell proliferation, the invasion and migration ability of the spheroid cells compared to their respective parental cells were carried out. Spheroids of both cell types demonstrated higher cell proliferation and migration abilities relative to their parental cells when subjected to the proliferation, migration and wound healing assays. Spheroids cells also exhibited higher invasion ability than parental cells. These results are in agreement with previous reports (Ludwig, Tse et al., 2013, You, Li et al., 2018), indicating that spheroids enriched with CSCs could adhere, migrate and invade more effectively than their parental counterparts. However, a study on cell motility of breast cancer spheroids showed that spheroid cells had a reduced cell invasion compared to the parental cells. We speculate that this contradiction may be explained by the differences in the methods used to enrich for the CSCs population using the serum-free anchorage independent technique. Differences in culture conditions and on the basis of the assays being carried out may contributed to the contradicting results. Taken together, this data suggest that the phenotypic characteristics of the enriched CSCs may vary depending on the techniques used to isolate the CSCs. Importantly, the

higher cell invasion and migration properties of spheroid cells enriched with CSCs provide insights into breast cancer pathogenesis. Increased cell motility is one of the factors that lead to tumour cell invasion and metastasis in human malignancy.

CSCs have been postulated to be slow cycling cells and in quiescent state, both factors responsible for the preservation of self-renewal and critical factor for chemoresistance (De Francesco, Sotgia et al., 2018, Moore and Lyle, 2011). In agreement with this hypothesis, the only similar research investigating between breast CSCs and quiescence was performed by Pece and colleagues in 2007, where they observed a link between quiescence in breast tumours and spheroids enriched with CSCs. They demonstrated that spheroids derived from grade 3 primary breast tumours had a higher CSCs content than grade 1 tumours. This finding is consistent with our MDA-MD-231 spheroid cells which were found to be held at quiescent stage (Figure 4.42). The more aggressive characteristics of the cells could have undergone EMT, accounted for their higher enrichment of CSCs in the spheroids and displayed that mesenchymal-like phenotype (Mani et al., 2008). On the other hand, MCF-7 spheroid cells are actively dividing cells indicated by the increase of cells in the G2+M and S phase as compared to the parental cells (Figure 4.41). The proliferating cells in the spheroids are undergoing EMT, accounted for partial enrichment of CSCs in the spheroids, identified as epithelial-like CSCs (Liu, S. et al., 2014).

We hypothesised that our spheroids were heterogeneous in nature, where only a subset of populations were CSCs, and therefore further purifying for the CSCs populations from the spheroids based on CSCs markers are necessary to enrich for CSCs cells (Jiang, Yang et al., 2017). Our on-going research on sorted populations positive for breast CSCs markers revealed an increased cells population in the quiescent stage. It has also been suggested that cell lines which have been in passage for an extended time might have been altered biologically should be interpreted differently for their CSCs characteristics. Therefore, this somehow described that quiescent state of CSCs are more relevant when the cells are directly isolated from primary tumours, in which, they are more primitive in nature, instead of cultured cancer cell lines. Nevertheless, the quiescent state of CSCs still needs further investigations to support for this as work currently being done is on other types of cancers such as melanoma and pancreatic adenocarcinoma (Dembinski and Krauss, 2009, Roesch, Fukunaga-Kalabis et al., 2010), which might not reflect breast cancer entirely.

Epigenetic dysregulation in cancers is thought to play an important role in tumorigenesis (Cebola and Peinado, 2012, Malik and Brown, 2000). DNA hypermethylation has been extensively studied and has been linked to tumour formation, development and metastasis (Das and Singal, 2004, Florl, Steinhoff et al., 2004, Suzuki, Maruyama et al., 2012). However, DNA hypomethylation in cancer remains unclear, and only until recently this has been observed in the early stages of tumorigenesis (De Capoa, Musolino et al., 2003, Ehrlich, 2009).

In our study, we performed the global methylation assay on the parental and spheroid cell, and we observed a decrease in DNA methylation of the spheroid cells. DNA hypomethylation could have occurred in the spheroid cells that resulted in this reduction. This initial finding is the first reported on breast spheroid enriched CSCs, consistent with recent finding in ovarian cancer (Liao, Qian et al., 2014), in which they observed involvement of hypomethylation of certain genes and their increased stemness properties in the ovarian tumour-initiating cells. Whole genome promoter microarray on DNA methylation profiles also demonstrated higher hypomethylated regions in breast CSCs than non-breast CSCs, suggesting that cancer-linked hypomethylation is a new epigenetic regulator for development of CSCs (El Helou, Wicinski et al., 2014). In our study, we hypothesised that the three-dimensional microenvironment in which the cells in the spheroid reside in may enhance the cell-cell interactions and affect the epigenetic status of the cells, perhaps enhancing the enrichment of breast CSCs. Closer inspection at the cancer-linked hypomethylation of gene promoters regions specifically can help to provide insights into DNA hypomethylation and their association with CSCs mechanisms (Suzuki, Maruyama et al., 2012).

4.4.4 Current Limitations

There were several limitations identified in the present study. Firstly, the spheroids generated in our study were heterogeneous spheroids formed from cell aggregation and also enriched with self-renewal CSCs-like cells

instead derived from a clone of cells. The CSCs theory suggests that CSCs should be derived from a single clone of homogeneous cells, which is from a single cell entity (Lawson, Kessenbrock et al., 2018), which could be achieved experimentally by plating single cell per well of a culture plate. However, clonal expansion is limited in the context of cancer research as single cell expansion requires a long time and substantial funding to generate sufficient cells for any downstream analysis. Therefore, researchers have resolved to use cell aggregation technique to allow for spheroid formation to enrich for CSCs subpopulations. From our literature search, most of the spheroids generated for CSCs studies were actually aggregates of cells rather than expanded from a single clone of cell as summarised in Table 4.4. Therefore, it is becoming more evident that more research groups are using spheroid culture based on cell aggregation to investigate on the CSCs characteristics across different types of cancers (Ishiguro, Ohata et al., 2017). It has also been demonstrated that cells from clonal expansion could also resulted from cell aggregation due to the spontaneous locomotion of free-floating single cells and from the movement of plates during daily cell culture maintenance (Cui, Hartanto et al., 2016). As such, it is generally accepted that clonal and aggregates of cells might co-exist and resulted in a heterogeneous spheroid, which can be enriched with CSCs-like cells in a serum-free condition.

Table 4.4 A summary of CSCs models of different culture techniques

Experimental Design						
Cancer type	**Implication**	**Single**	**Low**	**High**	**Details**	**Reference**
Prostate (HPET)	CSCs model			/	2 $X10^4$ cells/mL in SFM	(Gu, Yuan et al., 2007)
Breast (MCF7 and SUM159)	CSCs model		/		500 to 1000 cells/mL	(Li, Zhang et al., 2010)
Ovarian (OVCAR3 and SKOV3)	CSCs model			/	5 X 10^3 cells in SFM	(Nam, Lee et al., 2012)
Oral (SAS and OECM-1)	CSCs model			/	5 x 10^4 cells/dish	(Chen, Chang et al., 2012)
Cervical (HeLa, SiHa, Ca, Ski, C-4-1)	CSCs model			/	1 x 10^3 cells/mL in ULP	(Lopez, Poitevin et al., 2012)
Gastric (MKN-45)	CSCs model		/		100 cells/well in ULP	(Liu, Ma et al., 2013)
Breast (MDA-MB-231)	CSCs model			/	1 $X10^4$ cells in ULP	(Abboodi, 2014)
Breast tumours	CSCs model			/	2 $x10^3$ cells/well in ULP	(Klevebring, Rosin et al., 2014)
Breast (BT474, MDA-MB-231)	Spheroid model			/	1 x 10^4 cells in 1% methycellulose	(Tancioni, Miller et al., 2015)
Ovarian (OVCAR-3)	CSCs model			/	1 $x10^5$ cells/mL in ULP	(Vermeersch, Wang et al., 2015)

Reference
Single: Clonal cell expansion
Low: Low density cell aggregation (100 to 1000 cells)
High: High density cell aggregation (>1000 cells)
SFM: Serum-free medium
ULP: Ultra low attachment plate

4.5 Summary

The use of spheroid culture technique enriches a subpopulation of CSCs in both of the breast cancer subtypes. It appears that the spheroid cells generated were phenotypically different particularly in their CSCs-like characteristics. Nevertheless, both spheroid cells demonstrated increased in self-renewability, expression of stem cell- and CSCs-related markers, possessed higher cell proliferation, exhibited migration and wound healing abilities, and demonstrated a higher resistance against tested chemotherapy drugs in relative to their parental cells. However, higher proportions of CSCs subpopulations were enriched in MDA-MB-231 spheroid as compared to the MCF-7 spheroid, indicated by their increased ALDH activity. The spheroid models developed from the two most commonly used breast cancer cell lines were in line with that of CSCs isolated from different types of cancers including prostate, colon and melanoma.

CHAPTER 5

RESULTS AND DISCUSSION PART 2: QUALITY ASSESSMENT OF MIRNAS LIBRARIES OF ALL SAMPLES PREPARED FOR NEXT GENERATION SEQUENCING

5.0 Results and Discussion Part 2: Quality assessment of miRNAs libraries of all samples prepared for next generation sequencing

5.1 Introduction

Progress in the development of advanced gene expression studies is dedicated to the understanding of dynamic biological processes in cells and tissues. In the last decade, the discovery of new and previously undetected miRNAs as biomarkers has been made possible using this improved powerful platform of next generation sequencing (NGS) (Di Resta and Ferrari, 2018). NGS has been a new choice for transcriptomic analysis as it is more robust, has increased sensitivity, and higher accuracy. NGS applications have been increasingly widely used in many research areas which now span from agriculture to medical sciences (Ma, Gong et al., 2017, Shabani Azim, Houri et al., 2018). Consequently, due to the many advantages NGS can offer, this technology has slowly replaced microarrays for analysis of gene expressions. All NGS protocols follow strict workflows which are usually begin with RNA extraction, quantification and quality assessments, and lastly sequence data handling and pre-processing (Endrullat, Glokler et al., 2016). In this current chapter, NGS technology was used to uncover the miRNA expression profiles in two spheroid-enriched CSCs models, and thus pave a way for a better understanding of the targeted genes, annotations and pathways in managing breast cancer. Current standardisation procedures for NGS workflows are still scarce as the some of the same standards applied in clinical sequencing are not applicable for plant genome sequencing for example. Herein, in the present study, the standards involved in miRNA-NGS of cultured cells were strictly

followed and modified where applicable, which might be helpful for development for future work in NGS.

5.2 Experimental Design

The flow chart in Figure 5.1 outlines the overall steps for the preparation of the starting materials, construction of miRNAs libraries for next generation sequencing run and pre-processing the miRNA-NGS data.

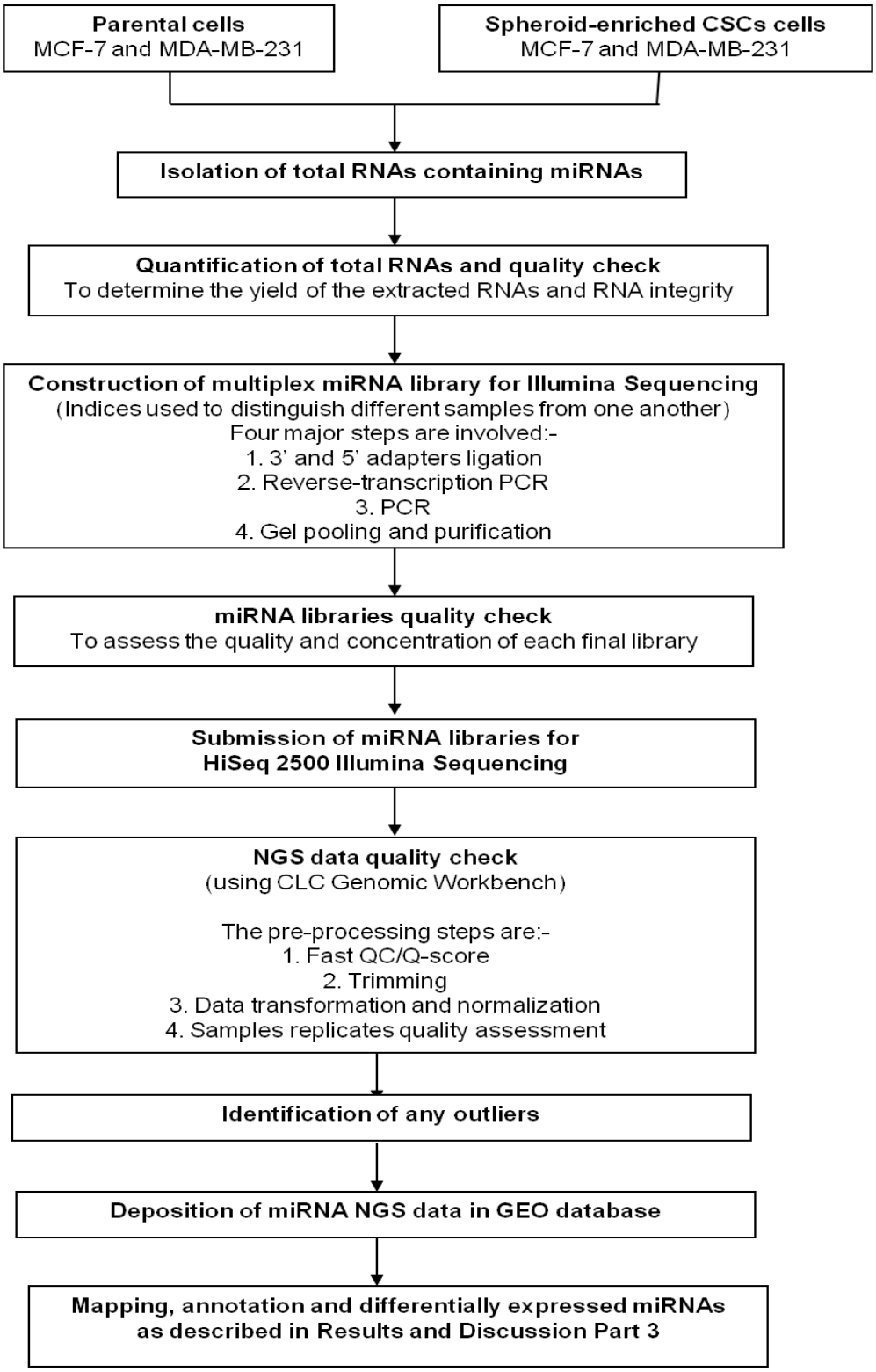


Figure 5.1. Experimental design of Part 2. Next generation sequencing workflow and quality assessment on the sequence data

5.3 Results

5.3.1 Concentration of Total RNAs and Quality Check

Total RNAs with retention of miRNAs were extracted from the parental and spheroids cells of MCF-7 and MDA-MB-231 breast cancer cell lines. The A260/A280 ratios of all the samples were above 1.8 indicating high purity and the A260/A230 ratios were above 1.7 indicating the samples were free from any impurities carry-over during the extraction process (Table 5.1). All the isolated RNAs were above 600 ng/µL indicating they were sufficient for the miRNA library preparation. Prior to RNA bioanalyzer integrity check, the RNAs were measured with Qubit RNA assay kit, for a more accurate measurement of the RNA yield (Table 5.2). The RNA concentration and purity of the extracted RNA samples were determined as only good quality RNA was suitable for subsequent experiment. Agilent Bioanalyzer using an RNA 6000 analysis Nano LabChip in Figure 5.2A showed that the all the samples had sharp distinction of both the 18S and 28S ribosomal RNA (rRNA) peaks. No shoulder bands appeared next to the two distinctive peaks indicating the RNA was of good quality. Electropherogram also showed that the intensity of the 28S band was approximately twice than of the 18S band, indicating intact RNA samples (Figure 5.2B). Small molecular weight RNA species with less than 200 base pairs were also observed on the electropherogram, appearing as intact and distinct, indicating that the small RNAs were successfully retained in the extracted RNA samples. RNA integrity number of all the RNA samples was >8 and all were suitable to be used in this study (Table 5.2).

Table 5.1. Nanodrop concentration and purity of the extracted RNA samples from the spheroids and parental breast cancer cell lines

Sample name	Biological replicate	A260/A280	A260/A230	Concentration (ng/µL)
MCF-7_P	1	1.98	1.95	630.30
MCF-7_P	2	2.00	1.92	624.00
MCF-7_P	3	2.4	2.24	1006.00
MCF-7_S	1	1.87	2.03	1940.00
MCF-7 _S	2	2.04	1.90	1672.00
MCF-7 _S	3	2.10	1.98	899.00
MDA _P	1	1.99	1.85	3114.00
MDA _P	2	2.05	2.20	1650.00
MDA _P	3	2.00	1.85	3114.00
MDA _S	1	2.01	1.95	3092.00
MDA _S	2	2.03	1.84	2340.00
MDA _S	3	2.01	2.00	1766.00

MCF-7_ P: MCF-7 parental cells
MCF-7 _S: MCF-7 spheroid cells
MDA_P: MDA-MB-231 parental cells
MDA_S: MDA-MB-231 spheroid cells
A260/A280 ratio > 1.80 and A260/A230 > 1.70, high quality RNA

Table 5.2. Qubit concentration and RIN indication of the extracted RNA samples from the spheroids and breast cancer cell lines

Sample name	Biological replicate	Qubit RNA concentration (ng/µL)	RIN indication
MCF-7_P	1	603.00	10.0
MCF-7_P	2	624.00	9.6
MCF-7_P	3	1955.00	10.0
MCF-7_S	1	4810.00	10.0
MCF-7 _S	2	2190.00	10.0
MCF-7 _S	3	1099.00	9.8
MDA _P	1	5700.00	9.9
MDA _P	2	3350.00	10.0
MDA _P	3	5700.00	9.9
MDA _S	1	4130.00	8.6
MDA _S	2	2390.00	8.3
MDA _S	3	2310.00	8.8

MCF-7_ P: MCF-7 parental cells
MCF-7 _S: MCF-7 spheroid cells
MDA_P: MDA-MB-231 parental cells
MDA_S: MDA-MB-231 spheroid cells
RIN>8, high quality RNA

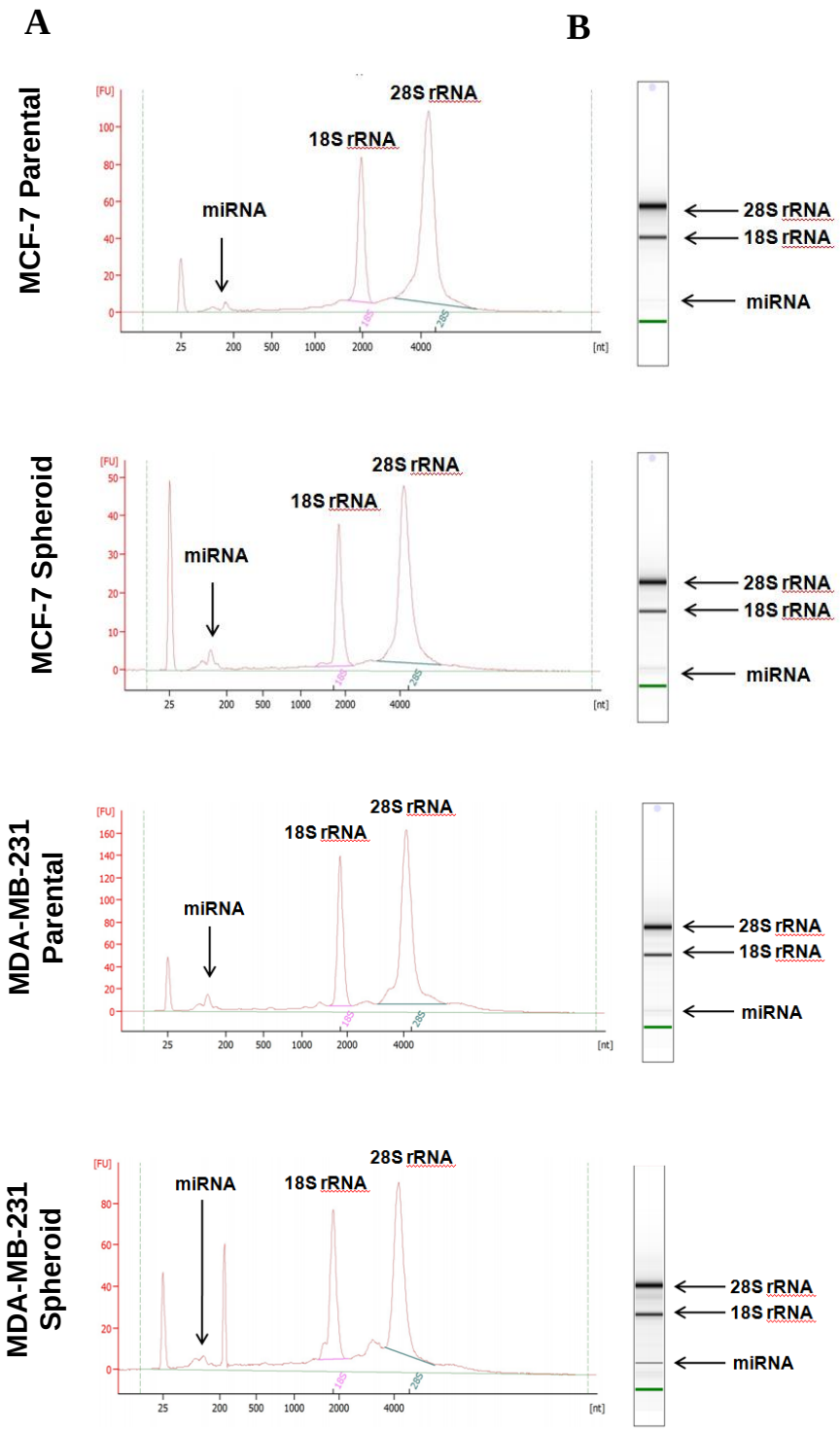


Figure 5.2. RNA integrity assessment by Agilent Bioanalyzer RNA 6000 LabChip. (A) Representative of electropherogram profiles of the good RNA peaks of the total RNA with retention of small RNA. (B) Gel electrophoresis of the electropherogram also revealed high quality RNA including the small RNA species. rRNA: ribosomal RNA.

5.3.2 miRNA Library Quality Check

To make use of the Illumina multiplexing capability for the miRNA-NGS, a total of 12 different index tags were used for all the samples. The indices comprised of six-base sequences as listed in Table 5.3 that helped to distinguish different samples from one another in a single lane of a flowcell during the NGS run. The DNA concentration for all the samples was determined using Qubit DNA (Table 5.4). All the samples had cDNA libraries of more than 10 ng/μL which were sufficient to be used for gel pooling and purification. Most importantly, the electropherogram profiles of the cDNA constructs from the miRNAs showed peaks of the expected miRNAs in the range of 145 bp to 160 bp (Figure 5.3). Common additional peaks were also observed indicating primer dimers (approximately 80 to 85 bp), adapter dimers (around 120 bp) and other regulatory small RNA molecules (piwi-RNA, tRNA, rRNA, and lincRNA). Individual libraries with unique indices from all the 12 samples were then pooled together into three libraries as listed in Table 5.5. The arrangement was made in accordance of their DNA concentration from the bioanalyzer run. Equal volumes of the libraries were loaded onto the PAGE gel (Figure 5.4), purified and ethanol precipitated into three final libraries. The quality for each of the library is an important determinant of the success of the sequencing run. The bioanalyzer pictured in Figure 5.5 showed that all the three libraries had only a single peak of the expected miRNAs at 146 bp, 147 bp, and 144 bp respectively, indicating the gel purification was successful. Bioanalyzer analysis also revealed that the three library concentration were > 10 nM (Table 5.6), which were more than enough based on the standard requirement for the library submission of at least a concentration of 2 nM.

Table 5.3. miRNA library preparation of all the samples and their unique indices

Sample name	Biological replicate	Index
MCF-7_P	1	RPI1
MCF-7_P	2	RPI2
MCF-7_P	3	RPI3
MCF-7_S	1	RPI4
MCF-7 _S	2	RPI5
MCF-7 _S	3	RPI6
MDA _P	1	RPI7
MDA _P	2	RPI8
MDA _P	3	RPI9
MDA _S	1	RPI10
MDA _S	2	RPI11
MDA _S	3	RPI12

MCF-7_ P: MCF-7 parental cells
MCF-7 _S: MCF-7 spheroid cells
MDA_P: MDA-MB-231 parental cells
MDA_S: MDA-MB-231 spheroid cells

Table 5.4. Qubit DNA concentration of the constructed cDNA miRNA library samples from the spheroids and parental breast cancer cell lines

Sample name	Biological replicate	Qubit DNA concentration (ng/µL)
MCF-7_P	1	24.5
MCF-7_P	2	19.6
MCF-7_P	3	12.6
MCF-7_S	1	25.8
MCF-7 _S	2	36.1
MCF-7 _S	3	18.7
MDA _P	1	44.8
MDA _P	2	23.3
MDA _P	3	11.9
MDA _S	1	14.7
MDA _S	2	18.4
MDA _S	3	18.5

MCF-7_ P: MCF-7 parental cells
MCF-7 _S: MCF-7 spheroid cells
MDA_P: MDA-MB-231 parental cells
MDA_S: MDA-MB-231 spheroid cells

Table 5.5. High sensitivity (HS)-DNA LabChip analysis of the constructed cDNA miRNA library samples from the spheroids and parental breast cancer cell lines

Sample name	Biological replicate	Peak detected (145 bp to 160 bp)	HS-DNA concentration (pg/µL)	Gel pooling arrangement
MCF-7_P	1	155	10.97	Pooled 1
MCF-7_P	2	156	495.09	Pooled 3
MCF-7_P	3	150	598.43	Pooled 3
MCF-7_S	1	154	35.05	Pooled 1
MCF-7 _S	2	157	72.04	Pooled 2
MCF-7 _S	3	155	466.56	Pooled 3
MDA _P	1	153	8.77	Pooled 1
MDA _P	2	151	22.89	Pooled 1
MDA _P	3	155	195.84	Pooled 2
MDA _S	1	155	165.06	Pooled 2
MDA _S	2	145	57.11	Pooled 2
MDA _S	3	148	321.89	Pooled 3

MCF-7_ P: MCF-7 parental cells
MCF-7 _S: MCF-7 spheroid cells
MDA_P: MDA-MB-231 parental cells
MDA_S: MDA-MB-231 spheroid cells

Gel pooling arrangement was divided into three groups based on the HS-DNA concentration as follows: -

Pooled 1: DNA concentration of 5 pg/µL to 50 pg/µL
Pooled 2: DNA concentration of 50 pg/µL to < 200 pg/µL
Pooled 3: DNA concentration of >200 pg/µL

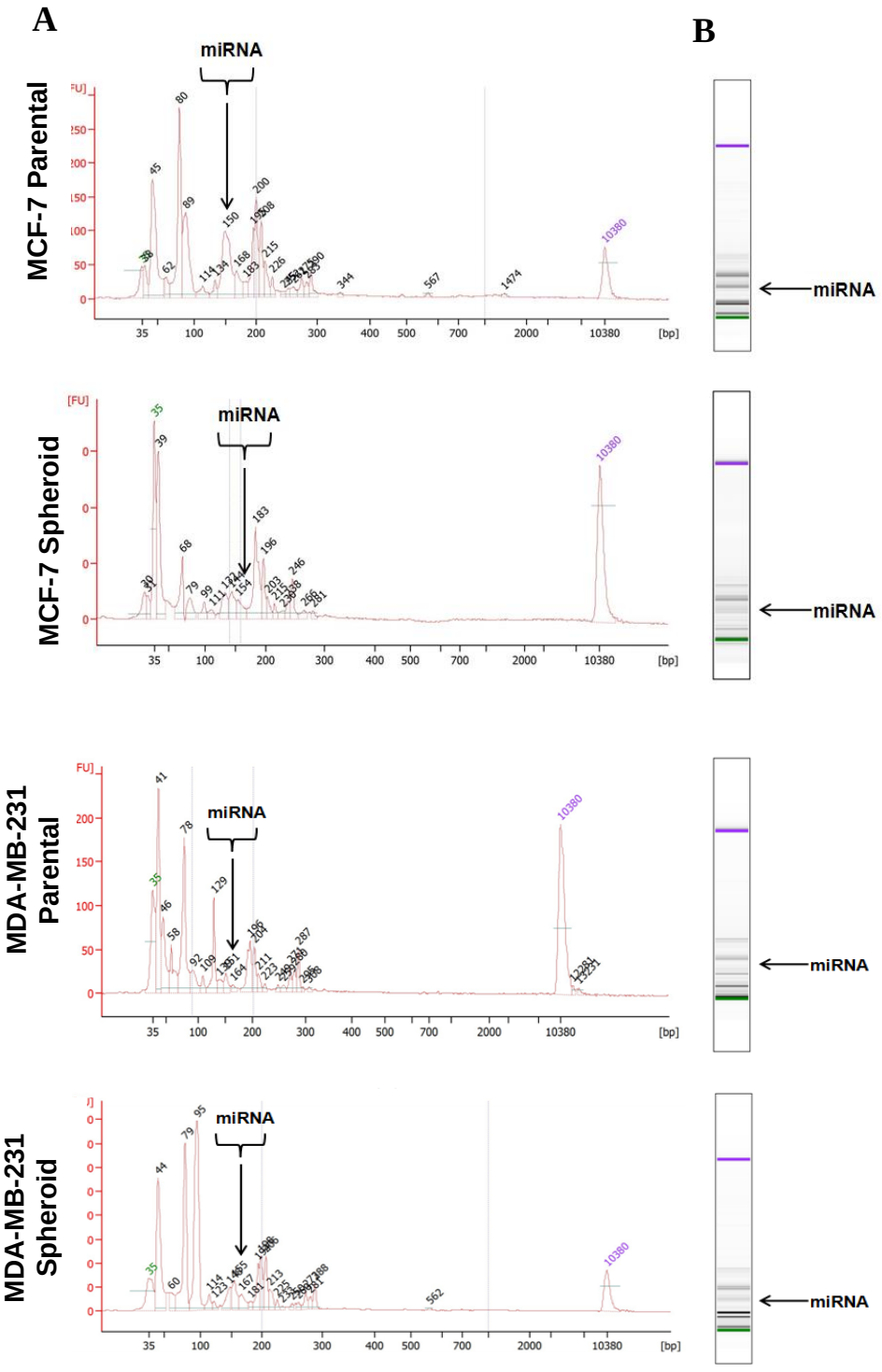


Figure 5.3. High sensitivity (HS) DNA LabChip of the constructed cDNA miRNA library samples from the spheroids and parental breast cancer cell lines. (A) Representative of electropherogram of the cDNA library after the final PCR step. (B) Gel electrophoresis of the electropherogram revealed the size of the miRNAs in the range of 145 to 160 bp. Other possibly bands seen are primers/primer dimers (<100 bp), amplified adapter dimer (120-125 bp), adapter concatamer (130-138 bp) and adapter-ligated-tRNA (>200 bp).

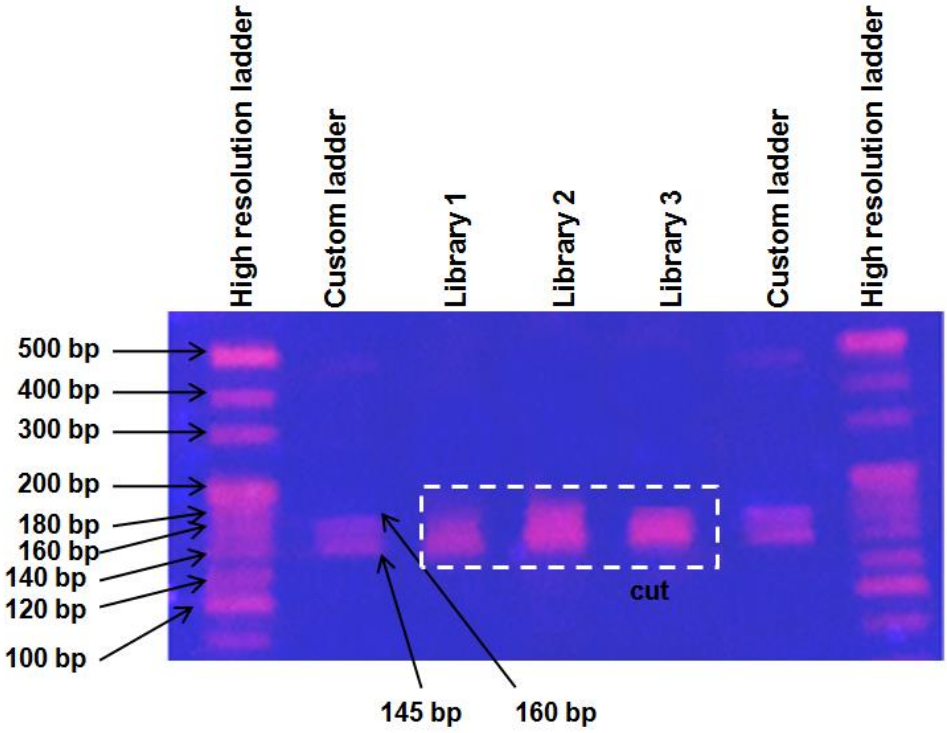


Figure 5.4. Small RNA library after PCR amplification separated on 6% PAGE gel. High resolution and custom ladder are shown on the left and right lanes. The libraries are indicated by the dotted white box. The gel fragments were excised by aligning the razor blade at the top of 160 bp band and at the bottom of the 145 bp of the Custom Ladder.

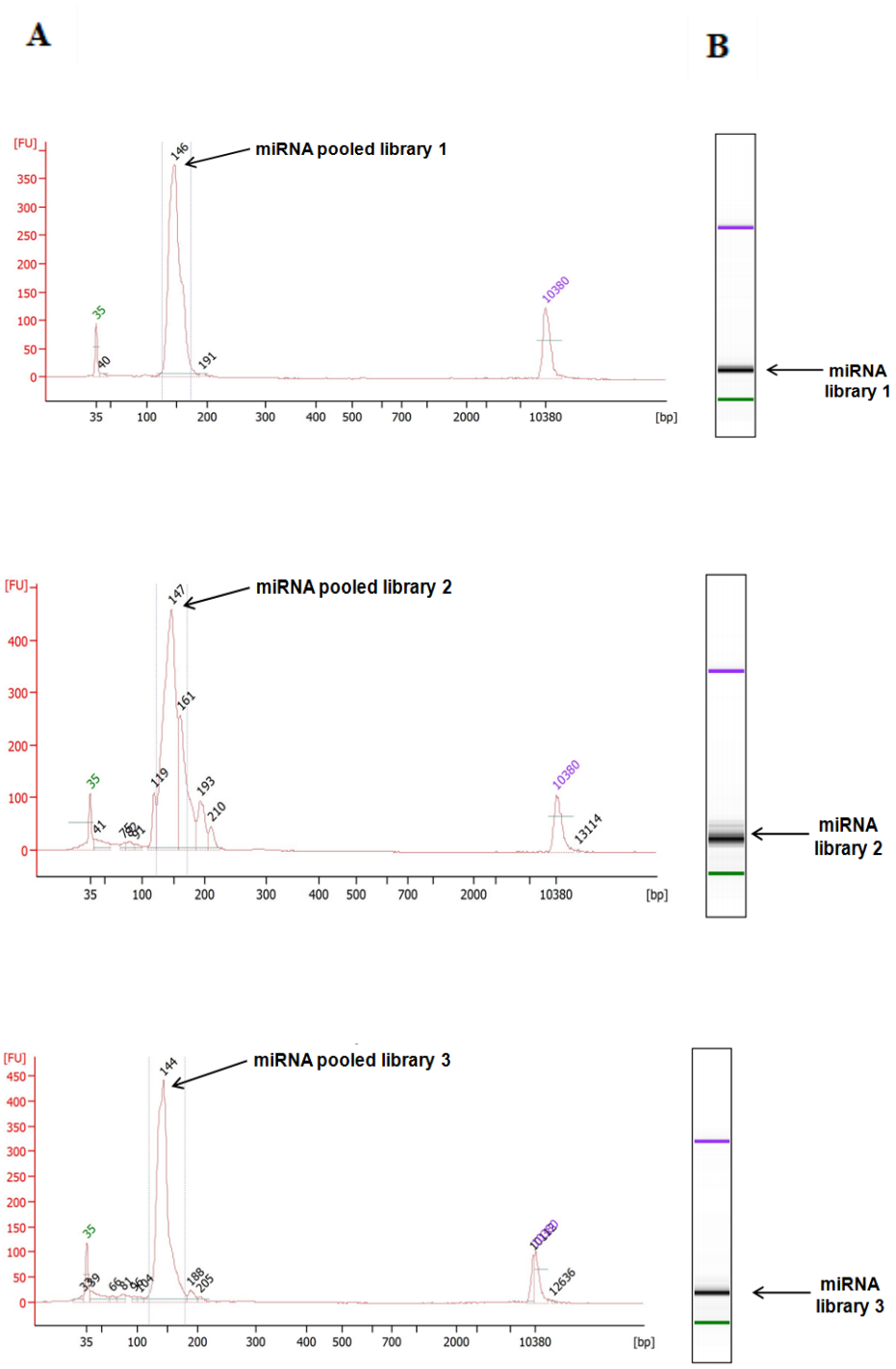


Figure 5.5. miRNA library size distribution on a High Sensitivity (HS) DNA Lab Chip (A) Electropherogram profiles shows a narrow distribution with a single library peak at the range of 145 to 150 bp (size of miRNAs) (B) Gel electrophoresis of the electropherogram revealed the correct size of the miRNAs of the different pooled libraries.

Table 5.6. High Sensitivity (HS) DNA LabChip of the constructed cDNA miRNA library samples after gel purification from the spheroids and parental breast cancer cell lines.

Sample name	Average size (bp)	Concentration (pg/µL)	Molarity (nmol)*
Pooled 1 miRNA Library	146	1441.54	14.88
Pooled 2 miRNA Library	147	2616.60	27.19
Pooled 3 miRNA Library	144	3102.46	32.84

Gel pooling arrangement was divided into three groups based on the HS-DNA concentration as follows: -

Pooled 1: DNA concentration of 5 pg/µL to 50 pg/µL
Pooled 2: DNA concentration of 50 pg/µL to <200 pg/µL
Pooled 3: DNA concentration of > 200 pg/µL

* All the libraries were diluted to 2 nM in total of 5 µL nuclease-free water before combined at equimolar ratios into a single multiplexed library for sequencing.

5.3.3 miRNA-NGS Quality-Based Data Pre-Processing

After obtaining the sequencing reads, the reads were deposited at the National Centre for Biotechnology Information Gene Expression Omnibus (GEO). For MCF-7 parental and spheroid samples, the GEO was approved with accession number GSE68246, which was publicly released on May 20th, 2016. On the other hand, for MDA-MB-231 parental and spheroids samples, the GEO was approved with accession number GSE75396, which was publicly released on November 25th, 2017.

Before the sequence reads were mapped to the reference genome, the reads of all the samples were subjected to FastQC quality control. The sequences were checked for the quality scores (Q-score) before and after trimming, and other basic statistics parameters available in the FastQC software (Table 5.7). When sequencing quality reaches Q-30, this is equivalent to the probability of an incorrect base call 1 in 1000 times (Table 5.8), whereby the base call accuracy is 99.9%. Therefore, Q-30 sequencing quality run means virtually all the reads are perfect, having zero errors and ambiguities. The reads were subjected to trim at Q-score 30 and to be discarded if the length is less than 17 bp or more than 25 bp after trimming. The majority of the trimmed reads for all the samples ranged from 20 to 26 nucleotides as shown in the read length distribution representative plots in Figure 5.6. The sequence length after trimming was found to be within the range of miRNAs length (Table 5.7). Based on the quality assessments, the sequencing reads in FASTQ format are in good condition.

Table 5.7. FastQC and basic statistics analysis of reads (before and after trimming) in each sample replicate

	Total sequences		Sequence quality (Q-score) *		Per sequence GC content		Adapter content		Sequence length	
Sample name	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**
MCF-7_P_1	4,946,582	3,484,952	23-38	37	53	51	12.97%	0.00%	50.0	23.6
MCF-7_P_2	25,403,527	14,172,793	33-37	37	50	48	6.02%	0.00%	50.0	23.6
MCF-7_P_3	2,912,607	1,640,491	33-37	37	52	51	6.45%	0.00%	50.0	23.4
MCF-7_S_1	5,547,764	2,825,065	27-37	37	55	54	1.64%	0.00%	50.0	23.2
MCF-7_S_2	6,821,442	4,184,171	27-37	37	55	54	2.07%	0.00%	50.0	26.0
MCF-7_S_3	6,832,659	4,190,684	33-37	37	57	58	3.07%	0.00%	50.0	26.0
MDA_P_1	1,292,680	895,445	27-37	37	52	50	14.97%	0.00%	50.0	22.0
MDA_P_2	1,814,513	1,406,056	33-37	37	51	48	5.46%	0.00%	50.0	22.9
MDA_P_3	18,989,485	10,865,502	33-37	37	51	48	3.71%	0.00%	50.0	23.3
MDA_S_1	8,403,201	5,636,651	33-37	37	51	41	12.72%	0.00%	50.0	20.7
MDA_S_2	9,045,488	7,066,458	33-37	37	48	40	2.23%	0.00%	50.0	20.7
MDA_S_3	9,068,719	7,081,696	33-37	37	48	40	2.29%	0.00%	50.0	22.5

MCF-7_ P: MCF-7 parental cells
MCF-7 _S: MCF-7 spheroid cells
MDA_P: MDA-MB-231 parental cells
MDA_S: MDA-MB-231 spheroid cells
Number 1, 2 and 3 indicates sample replicates.

* All the sequence obtained had Q-score >30, indicating good quality sequencing run.

Table 5.8. Quality scores (Phred score) and base calling accuracy (Ewing, Hillier et al., 1998)

Phred quality score	Probability of incorrect base call	Base call accuracy
10	1 in 10	90%
20	1 in 100	99%
30*	1 in 1,000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%

*****All the reads of all the samples had Q-score >30.

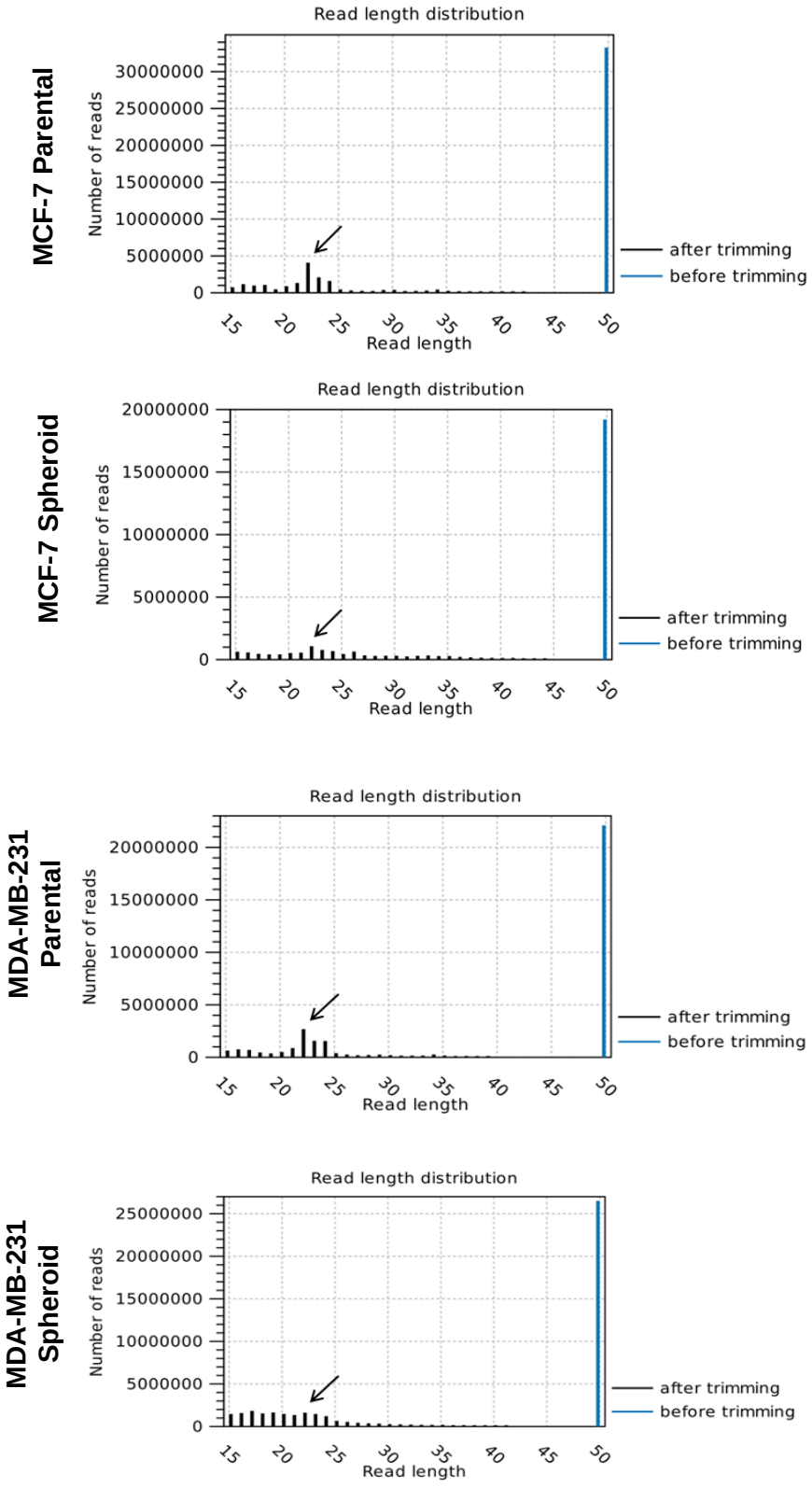


Figure 5.6. The representative read length distribution of miRNA reads. Arrows showing the average size of the reads of the different samples.

5.3.4 Samples Replicates Quality Assessment

All the sampled biological replicates were subjected to quality assessment analysis to determine the quality of each of the sample. It is important to remove any sources of variations between the biological samples which can occur during the miRNA library preparation, and during the sequencing run. Therefore, normalisation was performed to reduce this variation and to make the data comparable across all the samples. The miRNA sequencing transformed expression values were normalised using quantile method. Figure 5.7 showed the distribution of the expression values of the miRNA sequencing data using quantile normalisation. MiRNA pre-processing data analysis showed that that the size of the miRNAs of the different samples was in the acceptable range. The boxplots showed comparable distribution and no sample was identified as outlier after quantile normalisation. The quantile normalisation is the use of the empirical distributions of the sets of expression values for the samples to calculate a common target distribution. Following normalization, the group means for the two groups (parental versus spheroid cells) of both MCF-7 and MDA-MB-231 cells were also subjected to Scatter and Principal Component Analysis (PCA) plots. PCA plots on miRNA expression data from both cell types of spheroid and parental cells indicated the relative differential expression between groups (Figure 5.8A). Scatter plots of the group means for the two groups for each cell type between parental and spheroid cells indicate high mean expression values of the two datasets (Figure 5.8B). A hierarchical clustering of samples is a tree representation of their relative similarity. It was performed to depict the expression of the miRNAs between the spheroid and parental of the respective breast cancer cell lines

(Figure 5.9 and 5.10). Unsupervised hierarchical clustering analysis revealed that the spheroid cells were clearly separated from their parental cells. Expression of the differentially expressed miRNAs in parental and spheroid cells for both cell types was similar although some differences existed for specific miRNAs as indicated by the coloured horizontal lines.

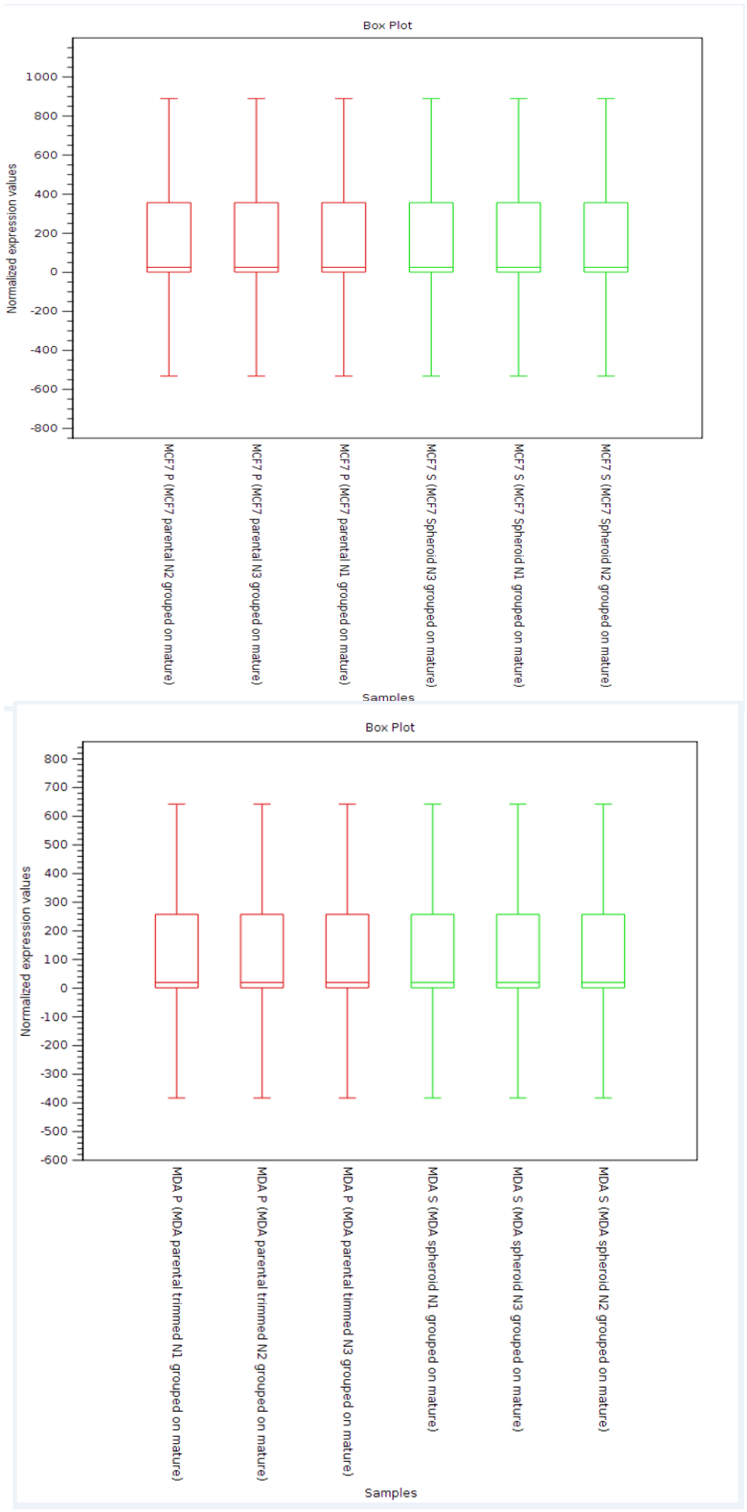


Figure 5.7. Expression data point distribution after quantile normalization by box and whisker plot. Cell types coloured by group, red: parental cells and green: spheroid cells. *x*-axis, sample name; *y*-axis, normalized expression values.

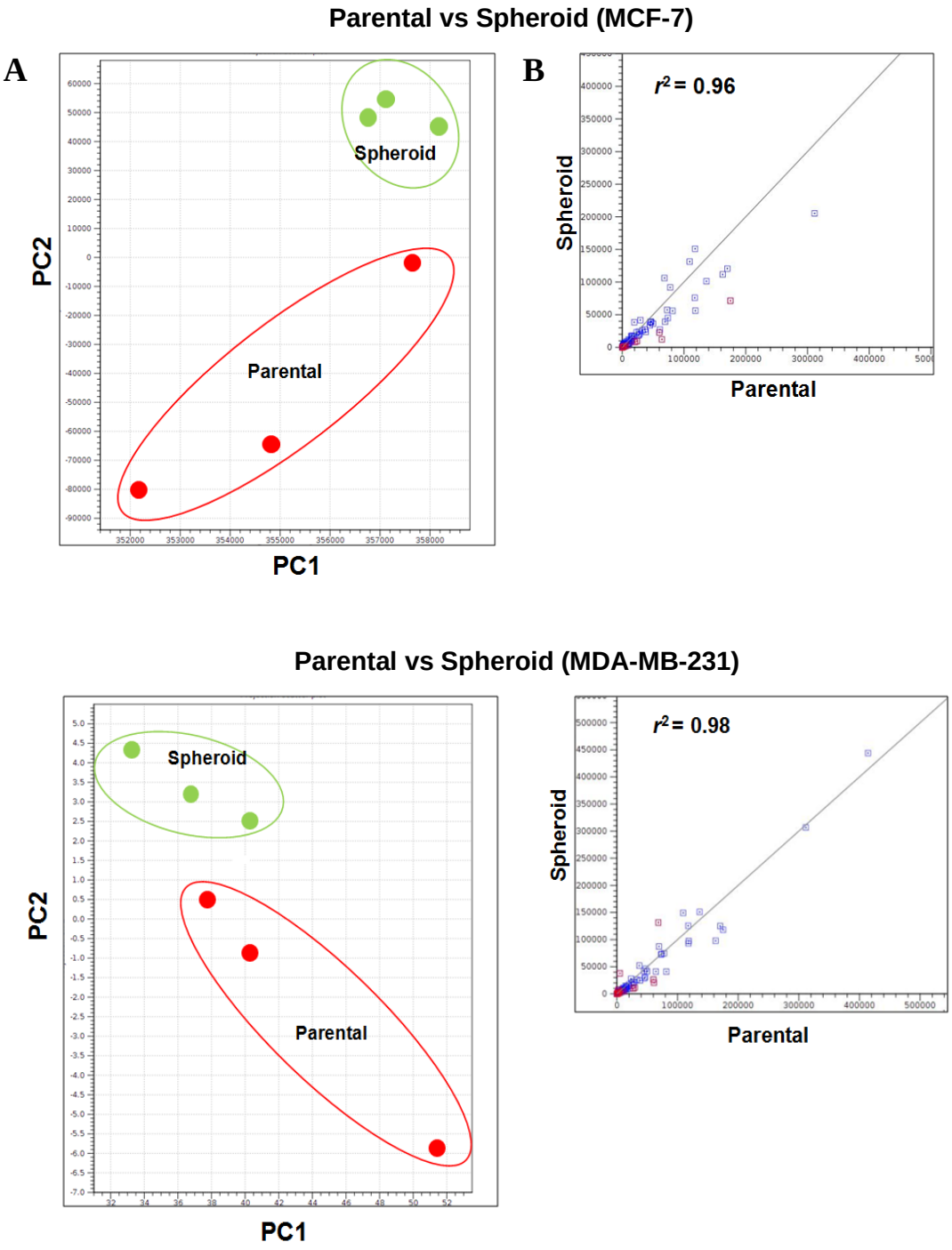


Figure 5.8. Principal component analysis (PCA) and scatter plots of all the normalised miRNA-NGS data. A) Each dot represents a biological replicate of a sample; different colour indicates different cell type. *x*-axis, first principal component (PC1); *y*-axis, second principal component (PC2). B) Scatter plots of the means miRNA-NGS data in each group. The correlation is shown. Top panel: MCF-7 parental vs spheroid; bottom panel: MDA-MB-231 parental vs spheroid.

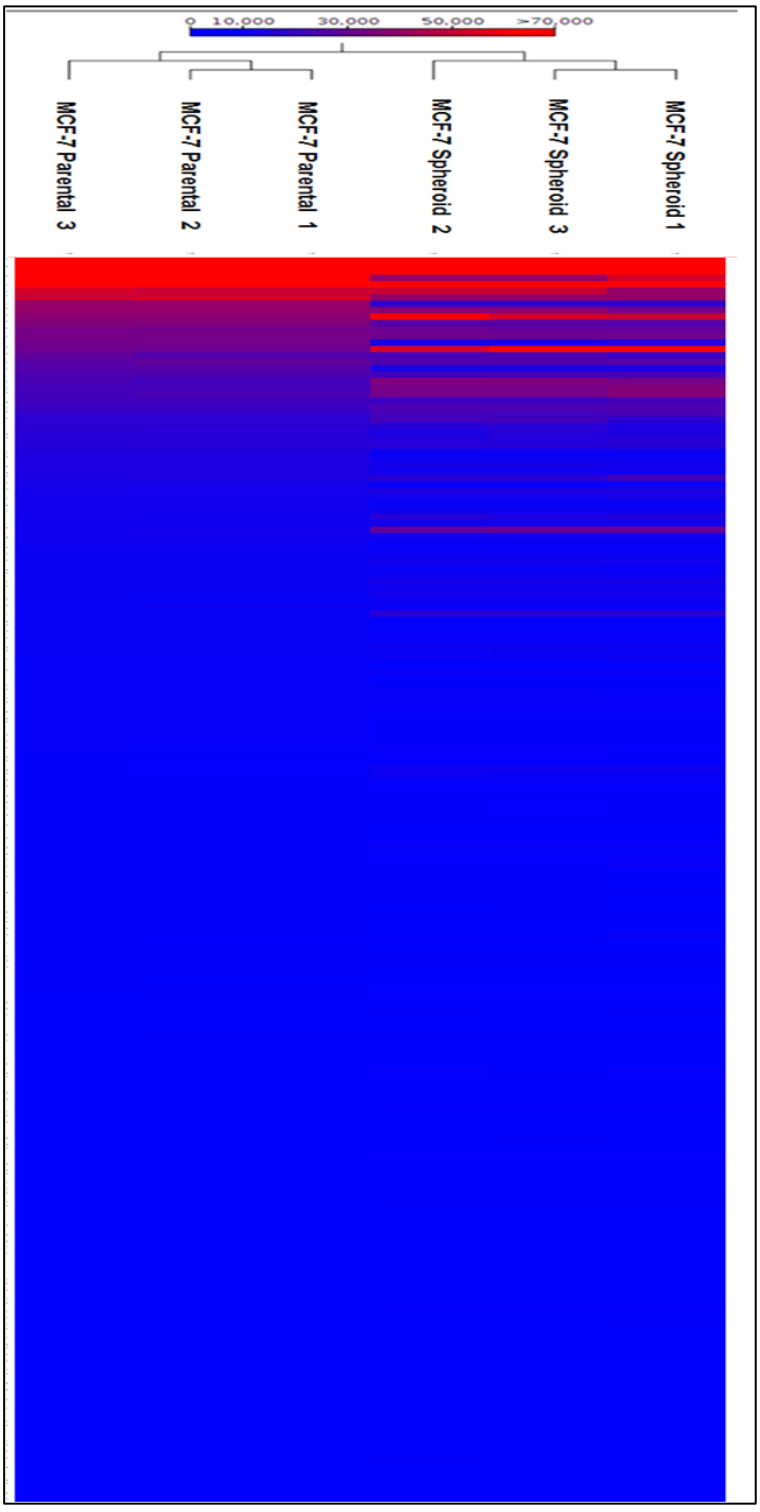


Figure 5.9. Heat map of the unsupervised hierarchical clustering (single linkage clustering, Pearson correlation coefficient) analysis of differentially expressed miRNAs and MCF-7 cell types. Data in the log2 scale averaged per biological replicate for each cell type. The names of the samples are listed at the top, the features are represented as horizontal lines, coloured according to the expression level.

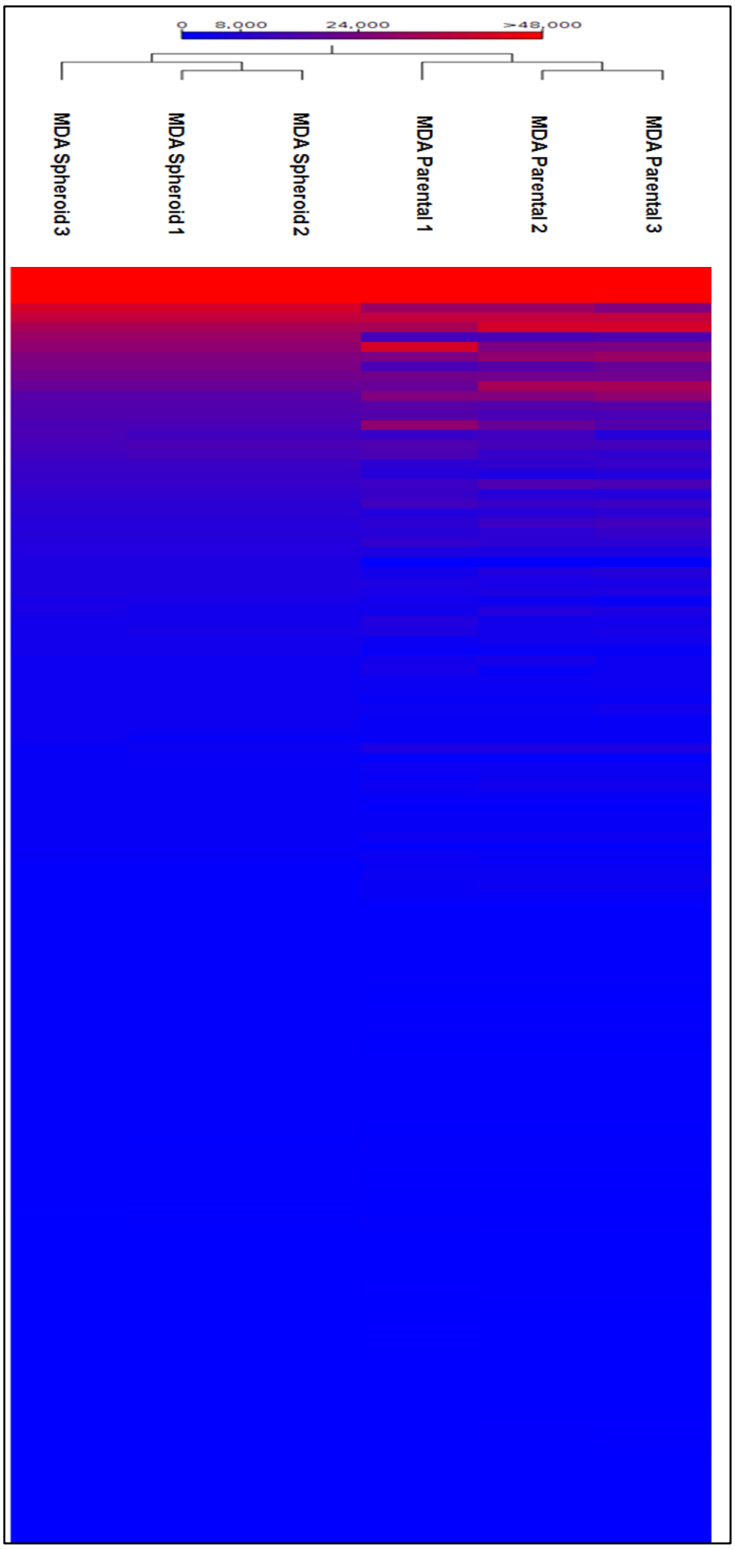


Figure 5.10. Heat map of the unsupervised hierarchical clustering (single linkage clustering, Pearson correlation coefficient) analysis of differentially expressed miRNAs and MDA-MB-231 cell types. Data in the log2 scale averaged per biological replicate for each cell type. The names of the samples are listed at the top, the features are represented as horizontal lines, coloured according to the expression level.

5.4 Discussion

5.4.1 Successful Extraction of Good Quality Total RNA

In this study, the quantity and quality of the total RNA of all the samples are important in order to obtain reliable results. RNA integrity number or commonly known as RIN number is an indication of RNA degradation given by predefined numerical numbers of 1 to 10, in which 10 is the highest quality of RNA obtained and 1 refers to the most degraded RNA (Schroeder, Mueller et al., 2006). It is largely accepted that RIN more than 8 indicate high quality RNA with minimum degradation that are commonly required for next generation sequencing or gene expression studies. The numbers are predicted using an algorithm that calculates the ratio between the large and small subunit of ribosomal RNAs (rRNAs). Degradation of RNA which can occur during the extraction stage or prolonged storage may compromise the downstream results. In addition, presence of contaminants such as organic compounds (e.g. phenol) also affects the quantification of total RNA. However, it has previously argued that RIN cannot be a good RNA indicator for those studies using miRNAs. This is likely consequence of the nature of miRNAs of being short sequence and is also extremely stable against RNAses. It is only useful for studies involving ribosomal RNA (rRNA) as it is made up of two subunits, the 28S and 18S subunits. Studies have also demonstrated that RIN value is irrelevant for miRNA analysis in which the miRNA expression levels across the different tissue types were not affected by the degradation of the heat-induced RNA (Jung, Schaefer et al., 2010). Comparative studies of the different efficiency of miRNA extraction kits also revealed that the low RIN obtained across the samples were not an indication of RNA degradation (El-Khoury, Pierson et al.,

2016). To assess the yield of total RNA in all the samples, a nanodrop was used and then further confirmed using Qubit assays. Qubit RNA assays which is a fluorescence labelling-based assay provides a more accurate results for the measurement of total RNA even in the presence of frequent contaminants such as proteins, DNA, nucleotides, or detergents. Therefore, this assay would give a better indication of the concentration of total RNA as quantity of input material resulted in significant changes of the NGS transcripts produced (Sarkar, Parkin et al., 2008). In our study, we have showed that RIN is a good indicator of high quality total extracted RNA with retention of miRNAs. The electropherograms from the bioanalyzer run were able to capture the miRNAs peaks at less than 200 bp, revealed that the miRNAs were retained in the extracted total RNAs, though with no indication of the integrity of the miRNAs. With next generation sequencing becoming a rapidly evolving and relevant technology today, indicator of overall quality of miRNA for NGS analysis have yet to be clearly established. Therefore, in the present study, combination of nanodrop and Agilent labchip technology should ideally represent an appropriate indication for the quantity and quality of total RNA with miRNAs extracted from all the samples. The data presented in our study have created a benchmark reference for methodology development which can be used for future studies on NGS involving miRNAs.

5.4.2 Construction of miRNA Libraries and Data Sequencing for Next Generation Sequencing

Next generation sequencing (NGS) is a powerful tool that resulted in large amount of sequence transcripts of different samples in one single run (Levy and Myers, 2016). However, a successful NGS run is largely depending on the high quality of the library preparation. In the present study, the general core steps involved in preparation of miRNA libraries for NGS analysis were followed in accordance to the manufacturer's instructions. However, during the initial pilot study, input of total RNA of 1 µg as suggested in the kit protocol did not produce any quantifiable miRNA library peaks within the 145 to 160 bp. This could be attributed to the fact that miRNA library preparation is more prone to adapter dimer formation when the RNA input is low (Dard-Dascot, Naquin et al., 2018). We therefore increased the input of the total RNA to 3 µg at the ligation step followed by the standard protocol, and surprisingly, this modification resulted in suppression of adapter dimer formation and increase miRNA library yields (Head, Komori et al., 2014). The final step in miRNA library preparation involves pooling of miRNA libraries and followed by gel purification to select the desired library size. Over the years, different methods have been established to prepare NGS libraries without any gel purification or more commonly known as "gel-free" methods. Chemical modifications of the adapters have been developed to prevent the formation of adapter primer and thus enhancing the ligation step and overcoming some of the challenges pertaining to gel purification technique (Shore, Henderson et al., 2016). Alternatively, Ampure beads have also been used for the size selection of the desired RNA library, but it needs to be carefully performed as lost samples and

low yields have been reported as some of the major drawbacks (Dard-Dascot, Naquin et al., 2018). Gel-free workflow allows researchers to work with lower sample input, a shorter library preparation time, and lastly facilitates automation of the whole process. These have been developed due the drawbacks of gel purification methods such as biasness, adapter primer contamination, reproducibility variation, and problems with sample recovery (Shore, Henderson et al., 2016). Though is a time-consuming step, gel purification remains a common choice as the use of new methods poses extraordinarily high cost compared to the conventional technique. The decision to use gel purification method may have been the most appropriate option in this study, since this approach has also been used in other study for the size selection of miRNA libraries (Marcial-Quino, Gómez-Manzo et al., 2016). Q-score or also known as Phred quality score is defined as a property that is logarithmically related to the base calling error probabilities, where the value can span from 0 to 50 (Ewing, Hillier et al., 1998). The better quality of the reads corresponded to the higher quality values. Low Q-scores generally increase false-positive variants calls resulted in inaccurate sequencing data. In this study, all the reads were above Q-score 30 indicating good quality sequencing run.

Biological replicates are necessary to assess and increase the precision of the analysis of NGS results (Robasky, Lewis et al., 2014). It is therefore important to assess the variations between the biological samples so that the differences between the miRNA expression patterns are due to the inherent heterogeneity of the independent samples. In the present study, the box plot,

PCA plot and the hierarchical clustering were able to characterise and differentiate between parental and spheroid samples in spheroids and also showed a clear global variation between the samples. The plots did not identify any outliers in the samples; therefore, no replicate was removed. However, further study involving higher numbers of biological replicates would be necessary in order to increase the reliability and to demonstrate higher degree of confidence in establishing the miRNA expression profiles. Increasing biological replicates has been demonstrated to improve accuracy in elucidating the genes expression profiles as compared to increasing the sequencing depth (Liu, Zhou et al., 2014).

5.5 Summary

To summarise, this study has clearly shown the importance of the yield and integrity of the total RNA with retention of miRNAs extracted from all the samples in order to ensure a reliable next generation sequencing result. In addition, the total input of total RNA which was increased to 3 µg instead of the suggested 1 µg has proven to be a better concentration to produce good enough miRNA libraries that can be validated downstream. All the biological replicates were in good condition, passed the quality assessment and were furthered used in the downstream analysis to identify differentially expressed miRNAs as discussed in the latter part of this thesis.

CHAPTER 6

RESULTS AND DISCUSSION PART 3: BIOINFORMATICS ANALYSIS OF COMMONLY AND UNIQUELY EXPRESSED MIRNAS IN SPHEROID-ENRICHED CSCS MODELS OF THE TWO BREAST CANCER SUBTYPES

Part of this chapter has been published in:

Boo, L., Ho, W. Y., Ali, N. M., Yeap, S. K., Ky, H., Chan, K. G., Yin, W. F., Dilan, A. S., Liew, W. C., Tan, S. W., Ong, H. K., Cheong, S. K. (2016). MiRNA transcriptome profiling of spheroid-enriched cells with cancer stem cell properties in human breast MCF-7 cell line. International Journal of Biological Sciences, 12(4), 427– 445.

and

Boo, L., Ho, W. Y., Ali, N. M., Yeap, S. K., Ky, H., Chan, K. G., Yin, W. F., Dilan, A. S., Liew, W. C., Tan, S. W., Cheong, S. K., Ong, H. K. (2017). Phenotypic and microRNA transcriptomic profiling of the MDA-MB-231 spheroid-enriched CSCs with comparison of MCF-7 microRNA profiling dataset. PeerJ, *5*, e3551.

6.0 Results and Discussion Part 3: Bioinformatics analysis of commonly and uniquely expressed miRNAs in spheroid-enriched CSCs models of the two breast cancer subtypes

6.1 Introduction

Small RNA molecules, being 19-25 nucleotides in length, known also as microRNAs (miRNA), has been found to play an important role in regulating gene expressions in normal and also in various human diseases, including cancer (Iorio, Ferracin et al., 2005, Porkka, Pfeiffer et al., 2007, Stefanie Sassen, Eric A. Miska et al., 2008). These single stranded RNAs regulate the expression of protein-coding genes at the translational level, either transcriptionally or post-transcriptionally (Di Leva and Croce, 2013a). The deregulation of miRNAs expressions has been observed specifically located in cancer-associated genomic regions (Blenkiron, Goldstein et al., 2007, Calin, Dumitru et al., 2002, Calin, Sevignani et al., 2004, Peng, Dai et al., 2013, Pereira, Marques et al., 2010). Although there are already some miRNAs that have been reported to be aberrantly expressed in breast cancer (Iorio, Ferracin et al., 2005), no specific study has been reported on the miRNAs signatures of breast cancer stem cell-like model (in our case, spheroid cells) by using miRNA-NGS. In addition, miRNAs-based NGS expression profiles in cancers are more informative compared with mRNA microarray profiles, supported by the fact that miRNA genes are located at fragile sites that are altered in human cancers. Using Ilumina-NGS, the present study was aimed to generate differentially expressed miRNAs of the spheroid and parental cells that could play a role in the progression of breast cancer specifically on self-renewal and drug resistance capacity of the spheroid-enriched CSCs. In this study, we

identified the spheroid-specific miRNA expression profiles followed by validation with qRT-PCR. We then performed functional annotation for the predicted target genes of these differentially expressed miRNAs with GO and KEGG analyses. Taken together, the data obtained in this study will be useful to improve the current understanding on their roles of miRNAs in the tumorigenesis of breast cancer.

Only a few studies have attempted to elucidate the differentially expressed miRNAs profiles using NGS. Thus far, the majority of the studies relied on microarray technology to generate the expression profiles of breast cancer CSCs or in breast cancer cell lines. Microarray expression technologies, though are widely used to generate whole genome expressions, they come with limitations (Shi, Liu et al., 2016, Wang and Xi, 2013). Extensive research has proven the inconsistency in results obtained using this system, which further questions the reliability of the results obtained. Contrary, NGS profiling with their higher accuracy, is a preferred choice for any type of profiling studies (Coenen-Stass, Magen et al., 2018).

Through NGS profiling, some miRNAs have been identified in the two most commonly used breast cancer cell lines, MCF-7 and MDA-MB-231. Only a handful of studies have attempted to identify the differentially expressed miRNAs in spheroid cells enriched CSCs models. Additionally, no study that has been done to identify the commonly expressed miRNAs by overlapping the cluster of miRNAs expressed in both spheroid subtypes models. Identification

of this subset of miRNAs is important as they could be the miRNAs modulating CSCs mechanisms. In the current study, when we compared the miRNA profiling of the MDA-MB-231 spheroid cells to the MCF-7 spheroid cells, we found a set of miRNAs exclusively expressed either in MDA-MB-231 or MCF-7 spheroid cells. A considerable amount of the differentially expressed miRNAs that were exclusively found was attributed to the fact that they were two distinct breast cancer subtypes.

6.2 Experimental Design

MiRNAs expression profiles were performed on three independent experiments for each group with TruSeq Small RNA Sample Preparation Kit (Illumina), followed by HiSeq run. The miRNAs expression profiles for four groups of samples are parental monolayer MCF-7, parental monolayer MDA-MB-231, MCF-7 spheroid-enriched CSCs cells and MDA-MB-231 spheroid-enriched CSCs cells. NGS data were then analysed with CLC Genomic Workbench 7.0 software for differentially expressed (DE) miRNAs, and DAVID software for target genes and pathway analysis. The DE miRNAs were then validated by real-time quantitative PCR. The flow chart in Figure 6.1 outlines the bioinformatics analysis, prediction of miRNA targets and functional analysis of target genes and lastly followed by miRNAs validation by real-time PCR.

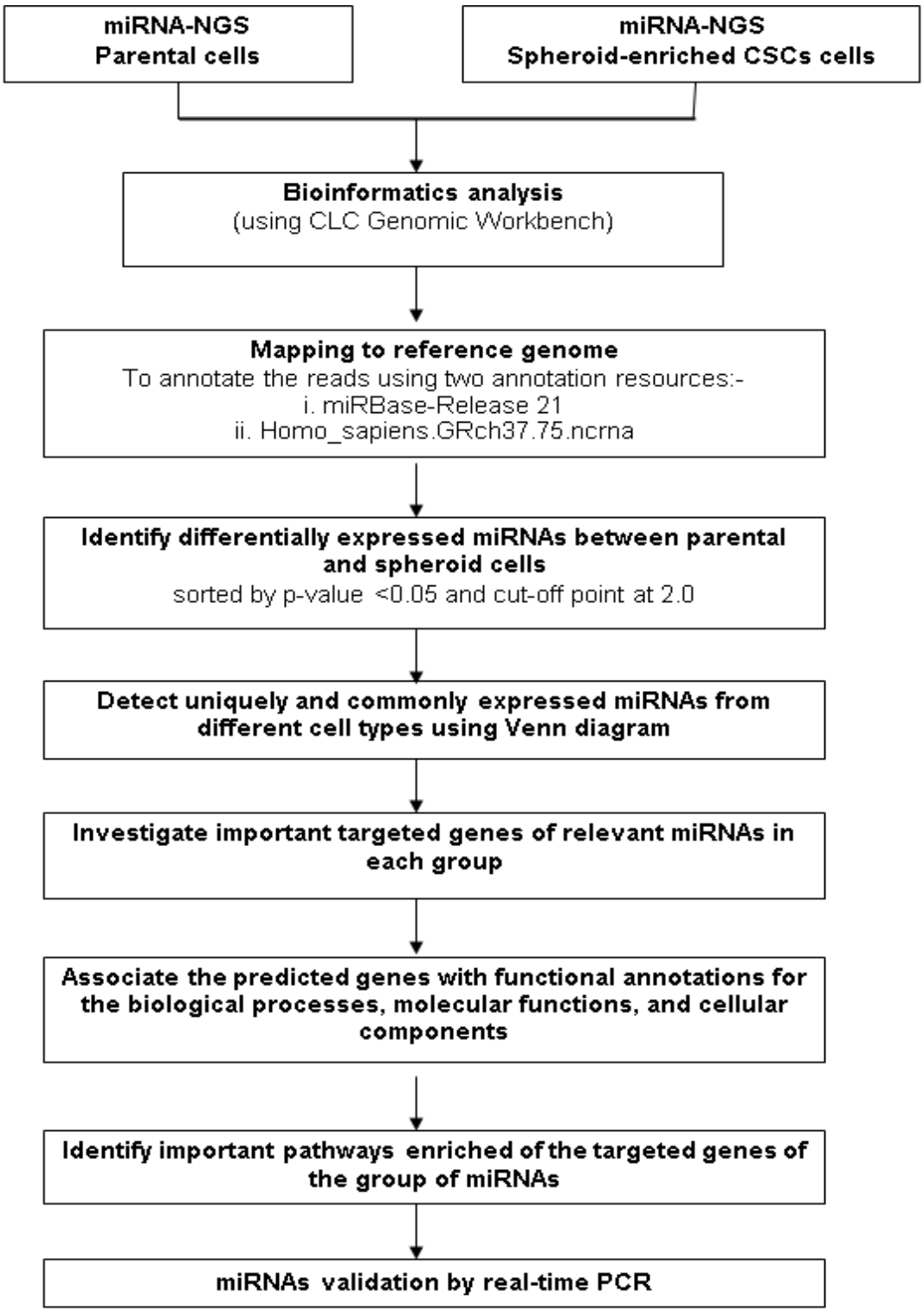


Figure 6.1. Experimental design of Part 3. Bioinformatics analysis workflow of miRNA-NGS and validation. The miRNA-NGS data was analysed via CLC Genomic Workbench version 7.0. The significantly differentially expressed miRNAs were determined and clustered into commonly and uniquely expressed miRNAs of the two cell types. Target genes belong to the miRNAs were elucidated together with their functional annotations in terms of gene ontologies and KEGG pathways.

6.3 Results

6.3.1 MiRNA Sequencing

The summary for small RNA library in parental and spheroids for both cell lines is listed in Table 6.1 and 6.2. The overall trend of this data confirms that the value obtained from the NGS was reliable. MiR-10a, miR-21 and miR-10b were observed to be the highest expression values though with different expression values in all the parental and spheroid samples for both cell types. These miRNAs were commonly found in breast cancer responsible for tumour growth, migration and invasion, response to tamoxifen drug, oncogenic properties, and progression (Fix LN, Shah M et al., 2010, Hoppe, Achinger-Kawecka et al., 2013, Jin, Wessely et al., 2013, Yan, Wu et al., 2011). Both miR-10a and miR-10b genes located in HOX clusters, a family of genes that code for transcription factors that were both involved in the transcription factors responsible in embryo development and also in oncogenesis (Lund, 2009, Ma, 2010). MiR-10b in particular, is known as an oncogene in breast cancer responsible for tumorigenesis promoting migration, invasion and metastasis, while very little is known about miR-10a in breast cancer. MiR-10a, on the other hand, is recognized by its higher expression than that of miR-10b in breast cancer cell lines (Ma, 2010). This is consistent with our NGS data analysis where we also observed the relatively higher expression of miR-10a compared to miR-10b in both MCF-7 (miR-10a: 929530/299888 reads in parental/spheroid vs miR-10b: 478725/168371 reads in parental/spheroid) and MDA-MB-231 (miR-10a: 65206/277306 reads in parental/spheroid vs miR-10b: 45231/128203 reads in parental/spheroid) cell lines.

Table 6.1. The top ten expressed miRNAs and their expression values observed in MCF-7 parental and spheroid cells.

MCF-7 parental cells		MCF-7 spheroid cells	
miRNA	**Expression values**	**miRNA**	**Expression values**
mir-10a	929530	mir-10a	299888
mir-21	847796	mir-21	270940
mir-27b	624659	mir-10b	168371
mir-92a-1//mir-92a-2	538802	mir-27b	82635
mir-10b	478725	mir-148a	39525
let-7a-1//let-7a-2//let-7a-3	265802	mir-26a-1//mir-26a-2	38182
mir-191	262456	mir-92a-1//mir-92a-2	34525
let-7i	184942	mir-191	33372
mir-99b	184251	let-7a-1//let-7a-2//let-7a-3	29609
mir-26a-1//mir-26a-2	176901	let-7i	26308

Table 6.2. The top ten expressed miRNAs and their expression values observed in MDA-MB-231 parental and spheroid cells.

MDA-MB-231 parental cells		MDA-MB-231 spheroid cells	
miRNA	**Expression values**	**miRNA**	**Expression values**
mir-21	108885	mir-10a	277306
mir-10a	65206	mir-21	268854
mir-10b	45231	mir-10b	128203
mir-27b	34934	mir-27b	88754
let-7a-1//let-7a-2//let-7a-3	21125	mir-26a-1//mir-26a-2	58206
mir-22	20043	mir-22	53336
mir-30a	15374	let-7a-1//let-7a-2//let-7a-3	44645
mir-191	13211	mir-30d	41313
let-7f-1//let-7f-2	12615	mir-92a-1//mir-92a-2	41084
mir-26a-1//mir-26a-2	11689	mir-486//mir-486-2	34829

6.3.2 Mapping to Reference Genome and Differentially Expressed miRNAs

The completely matched reads were annotated using two annotation resources; i) miRBase-Release 21 and ii) Homosapiens.GRCh37.75. ncrna. The summary for the miRNA distribution for all the samples is displayed in Figure 6.2. The majority of the miRNAs were successfully matched to the miRBase database for parental cells, with 81.57 ± 3.81% and 89.37 ± 4.45% match in MCF-7 parental and MDA-MB-231 parental cells respectively. The remaining 18.43 ± 3.81% reads and 10.63 ± 4.45% reads were aligned to the non-coding RNAs database for the same cell types. On the other hand, the mapping of miRNAs to miRBase database found in spheroid cells were slightly lower compared to the parental cells, at 60.43 ± 4.56% and 70.10 ± 3.90% in MCF-7 and MDA-MB-231 cells respectively. Higher alignment to the non-coding RNA database was found in MCF-7 spheroid (39.57 ± 4.56%) and MDA-MB-231 spheroid (29.90 ± 3.90%) relative to the parental cell types. The non-coding RNAs found across all the samples suggest the presence of other small RNA molecules such as tRNAs (transfer RNAs), rRNAs (ribosomal RNAs), snoRNAs (small nucleolar RNAs), siRNAs (small interfering RNAs), piRNAs (piwi-interacting RNAs) and lncRNAs (long non-coding RNAs) which could be detected using NGS beside the dominant miRNAs population. In an effort to identify significant differences in miRNA expression between the parental and spheroid cells for MCF-7 and MDA-MB-231 cell type, differentially expression analysis was performed. Volcano plot analysis is a visual approach to inspect the results of the statistical analysis. It shows the relationship between the *p*-values of a statistical test and the magnitude of the difference in

expression values of the samples. The differentially expressed miRNAs in spheroid cells relative to the parental cells were sorted by *p*-value <0.05 and significant log2 fold change values with a cut-off point set at 2.0. Therefore, the differentially expressed miRNAs which *p*-values <0.05 appeared as coloured dots above the green horizontal threshold line in the Volcano plot (Figure 6.3 and 6.4). Using 2-fold expression as the cut-off point, 25 up-regulated and 97 down-regulated differentially expressed miRNAs were identified in MCF-7 spheroid cells compared to their parental (red dots on the left and right of the green vertical threshold lines). On the other hand, 30 up-regulated and 36 down-regulated differentially expressed miRNAs were found in MDA-MB-231 spheroid in relative to the parental cells. Our data suggest that there were distinct miRNA expression profiles in spheroid relative to parental cells for both breast cancer cell lines. This reflects that the phenotypic behaviour and other distinctive features of spheroid enriched CSCs in MCF-7 and MDA-MB-231 are regulated by miRNAs. The significant differentially expressed miRNAs in MCF-7 and MDA-MB-231 spheroid when compared to their parental cells are listed in Table 6.3 and Table 6.4 respectively.

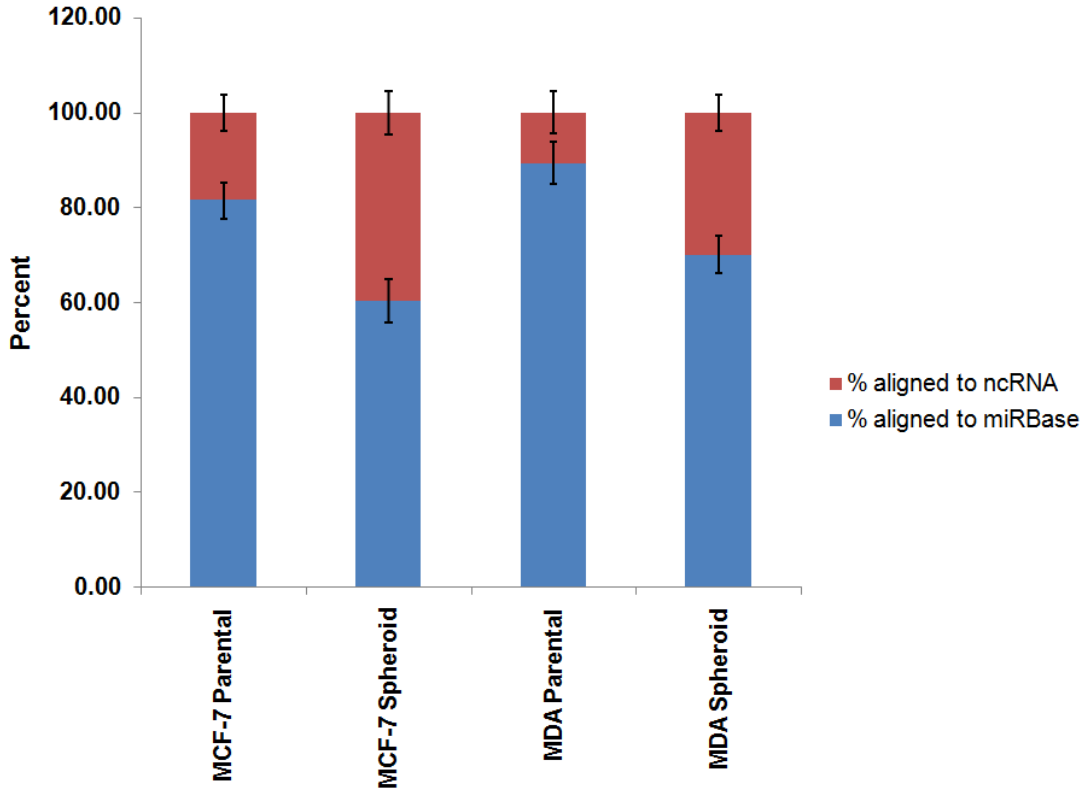


Figure 6.2. Read distribution of small RNA library sequencing results. The majority of miRNA species were aligned to the miRBase database. A smaller fraction of reads was mapped to the Homo sapiens GRCh38.ncRNA database.

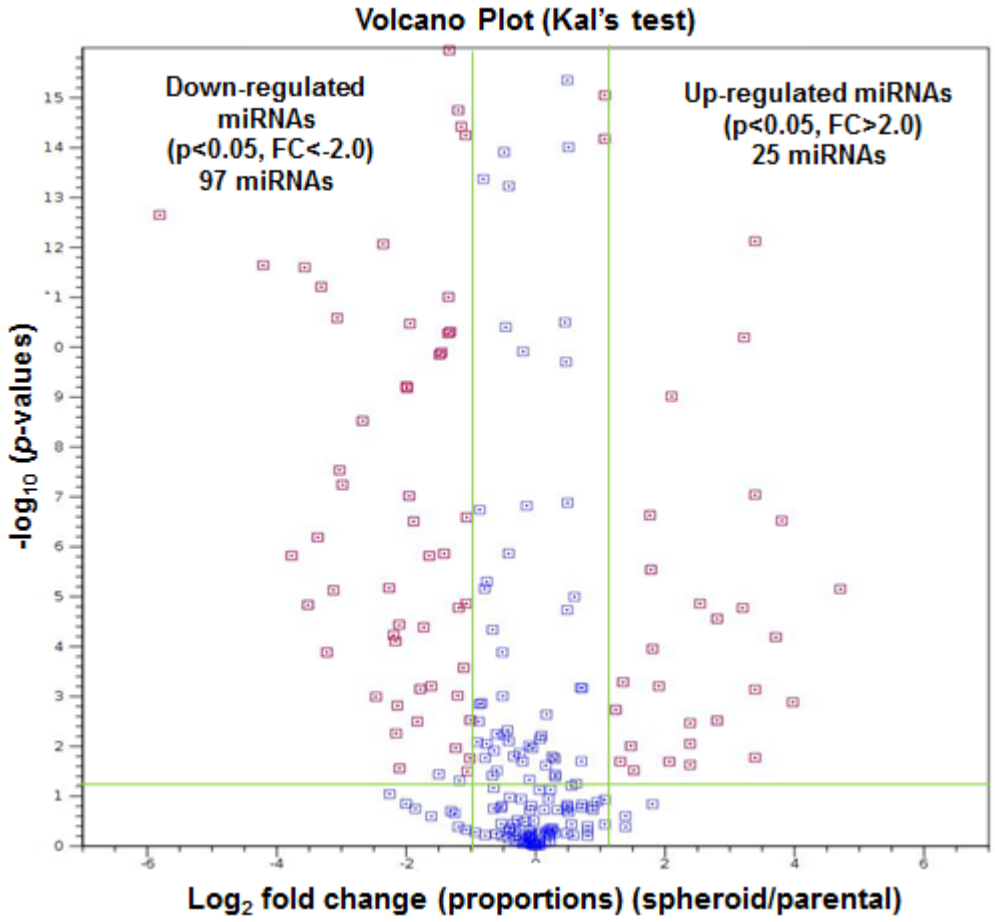


Figure 6.3. Volcano plot analysis of miRNA-NGS data of MCF-7 cell type. The analysis was performed using miRNA transcriptomics analysis in CLC Genomics Workbench. Volcano plot showing significantly (FC>2.0, p<0.05) differentially expressed miRNAs in MCF-7 spheroid relative to parental cells. Statistically significant miRNAs (p<0.05) are indicated above the green horizontal threshold line whereas miRNAs with significant log2 (fold change) values are indicated on the left and right of the green vertical threshold lines. Dots in blue represent miRNAs that do not have significant changes in expression, while dots in red on the left indicate miRNAs with significantly down-regulated expression and on the right indicates the miRNAs with significantly up-regulated expression. FC: fold change.

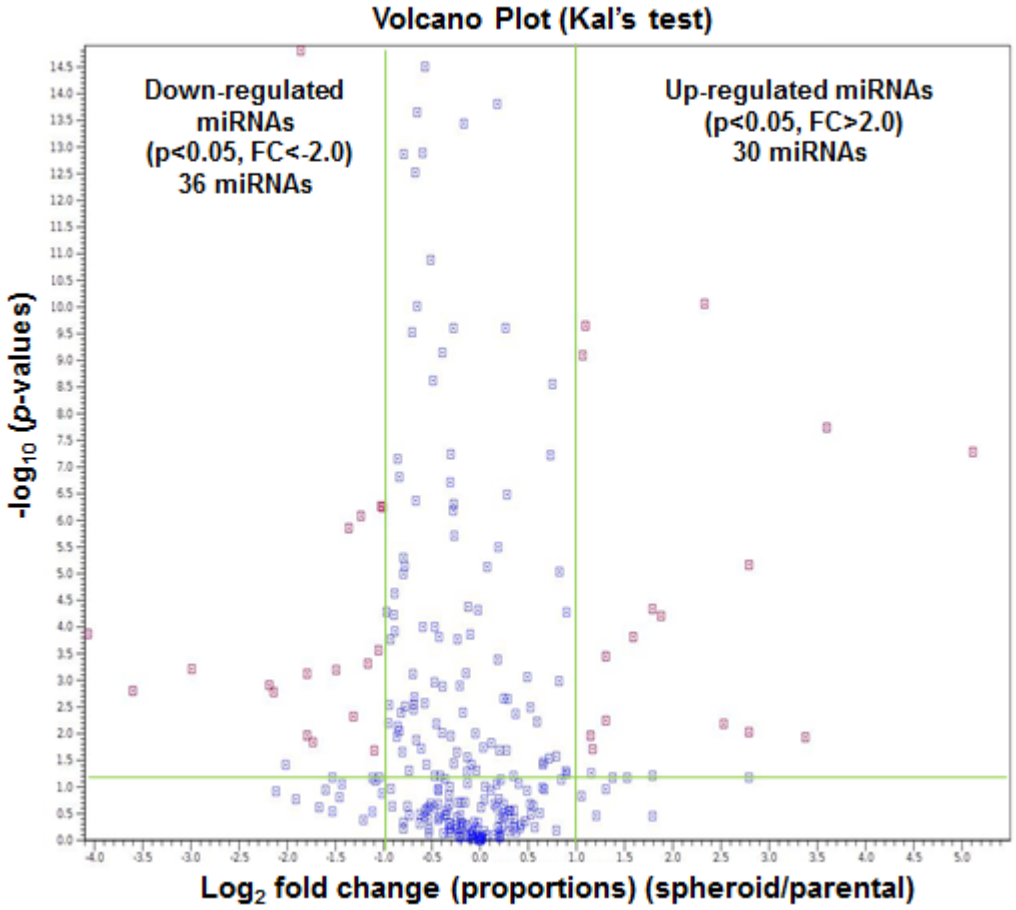


Figure 6.4. Volcano plot analysis of miRNA-NGS data of MDA-MB-231 cell type. The analysis was performed using miRNA transcriptomics analysis in CLC Genomics Workbench. Volcano plot showing significantly (FC>2.0, p<0.05) differentially expressed miRNAs in MDA-MB-231 spheroid relative to parental cells. Statistically significant miRNAs (p<0.05) are indicated above the green horizontal threshold line whereas miRNAs with significant log2 (fold change) values are indicated on the left and right of the green vertical threshold lines. Dots in blue represent miRNAs that do not have significant changes in expression, while dots in red on the left indicate miRNAs with significantly down-regulated expression and on the right indicates the miRNAs with significantly up-regulated expression. FC: fold change.

Table 6.3. miRNAs differentially expressed between spheroid and parental MCF-7 breast cancer cells

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Up-regulated miRNAs (n=25)	mir-4492	GGGGCTGGGCGCGCGCC	23.25	0.000E+00	0.000E+00
	mir-410	AATATAACACAGATGGCCTGT	15.00	0.000E+00	0.000E+00
	mir-4532	CCCCGGGGAGCCCGGCG	14.71	0.000E+00	0.000E+00
	mir-381	TATACAAGGGCAAGCTCTCTGT	10.00	0.000E+00	0.000E+00
	mir-127	TCGGATCCGTCTGAGCTTGGCT	8.66	0.000E+00	0.000E+00
	mir-411	TAGTAGACCGTATAGCGTACG	7.67	0.000E+00	0.000E+00
	mir-1246	AATGGATTTTTGGAGCAGG	6.24	0.000E+00	0.000E+00
	mir-409	GAATGTTGCTCGGTGAACCCCT	6.00	0.000E+00	0.000E+00
	mir-493	TTGTACATGGTAGGCTTTCATT	5.00	7.039E-06	1.727E-05
	mir-4508	GCGGGGCTGGGCGCGCG	4.80	0.000E+00	0.000E+00
	mir-143	TGAGATGAAGCACTGTAGCTC	4.54	0.000E+00	0.000E+00
	mir-126	CATTATTACTTTTGGTACGCG	4.12	0.000E+00	0.000E+00
	mir-1291	TGGCCCTGACTGAAGACCAGCAGT	3.75	0.000E+00	0.000E+00
	mir-370	GCCTGCTGGGGTGGAACCTGGT	3.00	1.309E-03	2.624E-03
	mir-136	CATCATCGTCTCAAATGAGTCT	2.67	2.988E-07	7.835E-07
	mir-145	GGATTCCTGGAAATACTGTTCT	2.50	6.495E-05	1.476E-04
	mir-211	TTCCCTTTGTCATCCTTCGCCT	2.50	6.495E-05	1.476E-04
	mir-378h	ACTGGACTTGGTGTCAGATGG	2.00	1.685E-02	2.978E-02
	mir-369	AGATCGACCGTGTTATATTCGC	2.00	1.685E-02	2.978E-02
	mir-99a	CAAGCTCGCTTCTATGGGTCTG	2.00	1.685E-02	2.978E-02
	mir-4485	TAACGGCCGCGGTACCCTAA	2.00	9.087E-08	2.475E-07
	mir-654	TATGTCTGCTGACCATCACCTT	2.00	7.507E-13	2.357E-12
	mir-499a	TTAAGACTTGCAGTGATGTTT	2.00	7.251E-04	1.528E-03
	mir-7641-1//mir-7641-2	TTGATCTCGGAAGCTAAGC	2.00	1.685E-02	2.978E-02
	mir-153-1//mir-153-2	TTGCATAGTCACAAAAGTGATC	2.00	1.685E-02	2.978E-02

Table 6.3 continued

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Down-regulated miRNAs (n=97)	mir-4448	GGCTCCTTGGTCTAGGGGTA	-60.00	1.467E-05	3.495E-05
	mir-221	ACCTGGCATACAATGTAGATTT	-57.25	0.000E+00	0.000E+00
	mir-27b	AGAGCTTAGCTGATTGGTGAAC	-56.89	0.000E+00	0.000E+00
	mir-125b-1	ACGGGTTAGGCTCTTGGGAGCT	-56.84	0.000E+00	0.000E+00
	mir-760	CGGCTCTGGGTCTGTGGGGA	-54.00	6.455E-07	1.666E-06
	mir-1296	TTAGGGCCCTGGCTCCATCTCC	-52.00	6.147E-12	1.882E-11
	mir-301a	GCTCTGACTTTATTGCACTACT	-49.00	1.307E-04	2.852E-04
	mir-365a	AGGGACTTTTGGGGGCAGATGTG	-45.67	7.470E-06	1.823E-05
	mir-30c-1	CTGGGAGAGGGTTGTTTACTCC	-43.86	2.603E-11	7.874E-11
	let-7f-1	CTATACAATCTATTGCCTTCCC	-43.84	0.000E+00	0.000E+00
	mir-4286	ACCCCACTCCTGGTACC	-43.71	0.000E+00	0.000E+00
	mir-671	TCCGGTTCTCAGGGCTCCACC	-42.92	0.000E+00	0.000E+00
	mir-26b	CCTGTTCTCCATTACTTGGCTC	-42.80	2.915E-08	8.073E-08
	mir-181a-2	ACCACTGACCGTTGACTGTACC	-41.78	0.000E+00	0.000E+00
	mir-615	GGGGGTCCCCGGTGCTCGGATC	-41.40	5.725E-08	1.577E-07
	mir-4454	GGATCCGAGTCACGGCACCA	-36.59	0.000E+00	0.000E+00
	mir-130b	ACTCTTTCCCTGTTGCACTAC	-33.70	0.000E+00	0.000E+00
	mir-193b	CGGGGTTTTGAGGGCGAGATGA	-33.38	2.985E-09	8.315E-09
	let-7b	CTATACAACCTACTGCCTTCCC	-32.44	0.000E+00	0.000E+00
	mir-26a-2	CCTATTCTTGATTACTTGTTTC	-30.67	2.021E-14	6.511E-14
	mir-155	CTCCTACATATTAGCATTAACA	-29.00	1.023E-03	2.128E-03
	mir-454	ACCCTATCAATATTGTCTCTGC	-26.81	8.558E-13	2.670E-12
	mir-423	AGCTCGGTCTGAGGCCCCTCAGT	-25.79	0.000E+00	0.000E+00
	mir-16-2	CCAATATTACTGTGCTGCTTTA	-25.14	6.743E-06	1.671E-05
	mir-191	GCTGCGCTTGGATTTCGTCCCC	-24.86	1.099E-14	3.589E-14
	mir-92b	TATTGCACTCGTCCCGGCCTCC	-24.07	0.000E+00	0.000E+00

Table 6.3 continued.

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Down-regulated miRNAs (n=97)	mir-628	ATGCTGACATATTTACTAGAGG	-24.00	5.889E-05	1.350E-04
	mir-424	CAAAACGTGAGGCGCTGCTAT	-23.88	3.331E-16	1.156E-15
	mir-3074	GATATCAGCTCAGTAGGCACCG	-23.50	7.820E-05	1.729E-04
	mir-30b	CTGGGAGGTGGATGTTTACTTC	-23.33	5.487E-03	1.049E-02
	mir-744	CTGTTGCCACTAACCTCAACCT	-23.00	1.528E-03	3.026E-03
	mir-100	CAAGCTTGTATCTATAGGTATG	-22.57	3.653E-05	8.494E-05
	mir-92b	AGGGACGGGACGCGGTGCAGTG	-22.50	2.774E-02	4.208E-02
	mir-99b	CACCCGTAGAACCGACCTTGCG	-22.06	0.000E+00	0.000E+00
	mir-7977	TTCCCAGCCAACGCACCA	-21.86	0.000E+00	0.000E+00
	mir-1468	CTCCGTTTGCCTGTTTCGCTG	-21.17	0.000E+00	0.000E+00
	mir-149	TCTGGCTCCGTGTCTTCACTCCC	-21.14	0.000E+00	0.000E+00
	mir-193a	TGGGTCTTTGCGGGCGAGATGA	-20.83	6.052E-10	1.715E-09
	mir-34a	CAATCAGCAAGTATACTGCCCT	-20.78	6.621E-10	1.865E-09
	mir-330	TCTCTGGGCCTGTGTCTTAGGC	-20.31	0.000E+00	0.000E+00
	mir-18a	ACTGCCCTAAGTGCTCCTTCTGG	-20.29	9.613E-08	2.604E-07
	mir-1307	ACTCGGCGTGGCGTCGGTCGTG	-20.28	0.000E+00	0.000E+00
	mir-152	AGGTTCTGTGATACACTCCGACT	-20.09	3.350E-11	1.001E-10
	mir-331	GCCCCTGGGCCTATCCTAGAA	-19.68	0.000E+00	0.000E+00
	mir-346	TGTCTGCCCGCATGCCTGCCTCT	-19.36	3.094E-07	8.067E-07
	mir-1260b	ATCCCACCACTGCCACCAT	-18.70	0.000E+00	0.000E+00
	mir-874	CGGCCCCACGCACCAGGGTAAGA	-18.60	3.180E-03	6.149E-03
	mir-484	TCAGGCTCAGTCCCCTCCCGAT	-18.36	0.000E+00	0.000E+00
	mir-148a	AAAGTTCTGAGACACTCCGACT	-18.07	0.000E+00	0.000E+00
	mir-365b	AGGGACTTTCAGGGGCAGCTGT	-18.00	7.099E-04	1.502E-03
	mir-125a	ACAGGTGAGGTTCTTGGGAGCC	-17.36	4.098E-05	9.484E-05

Table 6.3 continued.

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Down-regulated miRNAs (n=97)	mir-22	AGTTCTTCAGTGGCAAGCTTTA	-16.35	1.520E-06	3.843E-06
	mir-100	AACCCGTAGATCCGAACTTGTG	-16.02	0.000E+00	0.000E+00
	mir-338	AACAATATCCTGGTGCTGAGTG	-16.00	6.202E-04	1.332E-03
	mir-92a-1//mir-92a-2	TATTGCACTTGTCCCGGCCTGT	-15.61	0.000E+00	0.000E+00
	mir-125b-1//mir-125b-2	TCCCTGAGACCCTAACTTGTGA	-14.80	0.000E+00	0.000E+00
	mir-99b	CAAGCTCGTGTCTGTGGGTCCG	-14.66	1.413E-10	4.049E-10
	mir-30c-2	CTGGGAGAAGGCTGTTTACTCT	-14.35	1.279E-10	3.688E-10
	mir-224	CAAGTCACTAGTGGTTCCGTT	-14.23	0.000E+00	0.000E+00
	mir-423	TGAGGGGCAGAGAGCGAGACTTT	-14.18	0.000E+00	0.000E+00
	let-7c	TGAGGTAGTAGGTTGTATGGTT	-14.04	0.000E+00	0.000E+00
	mir-3613	ACAAAAAAAAAAGCCCAACCCTTC	-13.96	1.377E-06	3.518E-06
	mir-30d	CTTTCAGTCAGATGTTTGCTGC	-13.81	0.000E+00	0.000E+00
	mir-210	AGCCCCTGCCCACCGCACACTG	-13.37	5.289E-11	1.552E-10
	mir-1343	CTCCTGGGGCCCGCACTCTCGC	-13.34	9.862E-12	3.001E-11
	mir-1306	CCACCTCCCCTGCAAACGTCCA	-13.20	1.110E-16	3.882E-16
	mir-425	ATCGGGAATGTCGTGTCCGCCC	-13.06	4.889E-11	1.443E-10
	let-7d	CTATACGACCTGCTGCCTTTCT	-12.85	0.000E+00	0.000E+00
	mir-155	TTAATGCTAATCGTGATAGGGGT	-12.70	0.000E+00	0.000E+00
	mir-1260a	ATCCCACCTCTGCCACCA	-12.50	0.000E+00	0.000E+00
	mir-181a-1	ACCATCGACCGTTGATTGTACC	-12.48	0.000E+00	0.000E+00
	mir-196b	TCGACAGCACGACACTGCCTTC	-12.33	1.083E-02	1.984E-02
	mir-504	AGACCCTGGTCTGCACTCTATC	-12.24	0.000E+00	0.000E+00
	mir-135a-1	TATAGGGATTGGAGCCGTGGCG	-12.06	9.611E-04	2.016E-03
	mir-15b	TAGCAGCACATCATGGTTTACA	-12.05	0.000E+00	0.000E+00

Table 6.3 continued.

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Down-regulated miRNAs (n=97)	mir-331	CTAGGTATGGTCCCAGGGATCC	-12.01	1.776E-15	6.040E-15
	mir-1249	ACGCCCTTCCCCCCCTTCTTCA	-12.00	0.000E+00	0.000E+00
	mir-197	TTCACCACCTTCTCCACCCAGC	-11.96	0.000E+00	0.000E+00
	mir-877	GTAGAGGAGATGGCGCAGGG	-11.93	1.676E-05	3.973E-05
	mir-455	GCAGTCCATGGGCATATACAC	-11.71	0.000E+00	0.000E+00
	mir-361	TCCCCCAGGTGTGATTCTGATTT	-11.63	3.775E-15	1.275E-14
	mir-324	ACTGCCCCAGGTGCTGCTGG	-11.35	2.671E-04	5.801E-04
	mir-29a	TAGCACCATCTGAAATCGGTTA	-11.17	0.000E+00	0.000E+00
	mir-339	TCCCTGTCCTCCAGGAGCTCACG	-11.13	0.000E+00	0.000E+00
	mir-7-1//mir-7-2//mir-7-3	TGGAAGACTAGTGATTTTGTTGT	-11.08	5.662E-15	1.899E-14
	mir-22	AAGCTGCCAGTTGAAGAACTGT	-11.06	0.000E+00	0.000E+00
	mir-340	TCCGTCTCAGTTACTTTATAGC	-10.97	1.381E-05	3.304E-05
	mir-615	TCCGAGCCTGGGTCTCCCTCTT	-10.97	0.000E+00	0.000E+00
	mir-505	CGTCAACACTTGCTGGTTTCCT	-10.94	2.569E-07	6.773E-07
	let-7e	CTATACGGCCTCCTAGCTTTCC	-10.89	3.161E-02	4.736E-02
	mir-186	CAAAGAATTCTCCTTTTGGGCT	-10.71	0.000E+00	0.000E+00
	mir-15b	CGAATCATTATTTGCTGCTCTA	-10.60	0.000E+00	0.000E+00
	mir-98	CTATACAACTTACTACTTTCCC	-10.58	1.739E-02	3.050E-02
	mir-671	AGGAAGCCCTGGAGGGGCTGGAG	-10.53	2.995E-03	5.859E-03
	mir-148b	AAGTTCTGTTATACACTCAGGC	-10.47	0.000E+00	0.000E+00
	mir-3613	TGTTGTACTTTTTTTTTGTTC	-2.50	6.661E-15	2.219E-14
	mir-32	TATTGCACATTACTAAGTTGCA	-2.50	0.000E+00	0.000E+00
	mir-194-1//mir-194-2	TGTAACAGCAACTCCATGTGGA	-2.50	8.882E-16	3.041E-15
	mir-19a	TGTGCAAATCTATGCAAAACTGA	-2.39	0.000E+00	0.000E+00
	mir-542	TGTGACAGATTGATAACTGAAA	-2.37	0.000E+00	0.000E+00

Table 6.3 continued.

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
	mir-210	CTGTGCGTGTGACAGCGGCTGA	-2.29	0.000E+00	0.000E+00
	mir-545	TCAGTAAATGTTTATTAGATGA	-2.22	1.852E-03	3.651E-03
	mir-378d-2//mir-378d-1	ACTGGACTTGGAGTCAGAAA	-2.18	0.000E+00	0.000E+00
	mir-30e	TGTAAACATCCTTGACTGGAAG	-2.15	0.000E+00	0.000E+00
	mir-362	AATCCTTGGAACCTAGGTGTGAGT	-2.11	2.019E-02	3.398E-02
	mir-101-1//mir-101-2	TACAGTACTGTGATAACTGAA	-2.08	0.000E+00	0.000E+00
	mir-196a-2	CGGCAACAAGAAACTGCCTGAG	-2.05	5.187E-04	1.122E-03

Table 6.4 miRNAs differentially expressed between spheroid and parental MDA-MB-231 breast cancer cells

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Up-regulated miRNAs (n=30)	mir-411	TAGTAGACCGTATAGCGTACG	41.49	0.00E+00	0.00E+00
	mir-381	TATACAAGGGCAAGCTCTCTGT	34.58	5.19E-08	2.11E-07
	mir-127	TCGGATCCGTCTGAGCTTGGCT	33.84	0.00E+00	0.00E+00
	mir-143	TGAGATGAAGCACTGTAGCTC	17.73	0.00E+00	0.00E+00
	mir-133a-1//mir-133a-2	TTTGGTCCCCTTCAACCAGCTG	12.10	1.81E-08	7.41E-08
	mir-410	AATATAACACAGATGGCCTGT	10.37	1.17E-02	2.68E-02
	mir-145	GGATTCCTGGAAATACTGTTCT	10.37	1.17E-02	2.68E-02
	mir-126	TCGTACCGTGAGTAATAATGCG	7.06	0.00E+00	0.00E+00
	mir-1246	AATGGATTTTTGGAGCAGG	6.92	6.83E-06	2.37E-05
	mir-342	AGGGGTGCTATCTGTGATTGA	6.92	9.39E-03	2.23E-02
	mir-181a-1	ACCATCGACCGTTGATTGTACC	5.81	0.00E+00	0.00E+00
	mir-205	TCCTTCATTCCACCGGAGTCTG	5.76	6.59E-03	1.63E-02
	mir-210	AGCCCCTGCCCACCGCACACTG	5.68	0.00E+00	0.00E+00
	mir-125b-2	TCACAAGTCAGGCTCTTGGGAC	5.08	0.00E+00	0.00E+00
	mir-139	TCTACAGTGCACGTGTCTCCAGT	5.03	8.72E-11	3.90E-10
	mir-181a-2//mir-181a-1	AACATTCAACGCTGTCGGTGAGT	5.02	0.00E+00	0.00E+00
	mir-210	CTGTGCGTGTGACAGCGGCTGA	3.72	0.00E+00	0.00E+00
	mir-211	TTCCCTTTGTCATCCTTCGCCT	3.67	6.20E-05	1.98E-04
	mir-181b-1//mir-181b-2	AACATTCATTGCTGTCGGTGGGT	3.58	0.00E+00	0.00E+00
	mir-181c	AACCATCGACCGTTGAGTGGAC	3.46	4.62E-05	1.53E-04
	mir-126	CATTATTACTTTTGGTACGCG	3.43	0.00E+00	0.00E+00
	mir-199a-1//mir-199a-2//mir-199b	ACAGTAGTCTGCACATTGGTTA	3.31	0.00E+00	0.00E+00
	mir-1260a	ATCCCACCTCTGCCACCA	3.15	0.00E+00	0.00E+00
	mir-1260b	ATCCCACCACTGCCACCAT	2.53	0.00E+00	0.00E+00

Table 6.4 continued.

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
	mir-365b	AGGGACTTTCAGGGGCAGCTGT	2.47	3.55E-04	1.06E-03
	mir-1291	TGGCCCTGACTGAAGACCAGCAGT	2.47	5.66E-03	1.43E-02
	mir-885	TCCATTACACTACCCTGCCTCT	2.26	0.00E+00	0.00E+00
	mir-885	AGGCAGCGGGGTGTAGTGGATA	2.25	1.95E-02	4.34E-02
	mir-664b	TTCATTTGCCTCCCAGCCTACA	2.21	1.09E-02	2.54E-02
	mir-129-2	AAGCCCTTACCCCAAAAAGCAT	2.13	2.24E-10	9.83E-10
	mir-328	CTGGCCCTCTCTGCCCTTCCGT	2.09	8.01E-10	3.37E-09
	mir-30d	TGTAAACATCCCCGACTGGAAG	2.06	0.00E+00	0.00E+00
	mir-423	TGAGGGGCAGAGAGCGAGACTTT	2.06	0.00E+00	0.00E+00
Down-regulated miRNAs (n=36)	mir-3074	GATATCAGCTCAGTAGGCACCG	-16.77	1.34E-04	4.18E-04
	mir-190b	TGATATGTTTGATATTGGGTT	-12.15	1.57E-03	4.32E-03
	mir-769	CTGGGATCTCCGGGGTCTTGGTT	-7.95	6.15E-04	1.80E-03
	mir-7-1//mir-7-2//mir-7-3	TGGAAGACTAGTGATTTTGTTGT	-6.94	0.00E+00	0.00E+00
	mir-3613	TGTTGTACTTTTTTTTTTGTTC	-6.42	0.00E+00	0.00E+00
	mir-204	TTCCCTTTGTCATCCTATGCCT	-5.69	0.00E+00	0.00E+00
	mir-34c	AGGCAGTGTAGTTAGCTGATTGC	-4.55	1.24E-03	3.45E-03
	mir-503	TAGCAGCGGGAACAGTTCTGCAG	-4.41	1.65E-03	4.51E-03
	mir-33a	GTGCATTGTAGTTGCATTGCA	-3.63	1.55E-15	7.82E-15
	mir-200a	TAACACTGTCTGGTAACGATGT	-3.47	1.08E-02	2.54E-02
	mir-24-1	TGCCTACTGAGCTGATATCAGT	-3.47	7.49E-04	2.15E-03
	mir-1271	CTTGGCACCTAGCAAGCACTCA	-3.45	0.00E+00	0.00E+00
	mir-222	CTCAGTAGCCAGTGTAGATCCT	-3.40	5.10E-04	2.10E-01
	mir-802	CAGTAACAAAGATTCATCCTTGT	-3.33	1.44E-02	3.27E-02
	mir-574	TGAGTGTGTGTGTGTGAGTGTGT	-3.22	5.50E-07	2.27E-04
	mir-138-2//mir-138-1	AGCTGGTGTTGTGAATCAGGCCG	-2.96	0.00E+00	0.00E+00

Table 6.4 continued.

	miRNA	Feature ID	Fold change	Unadjusted p-values	FDR p-values correction
Down-regulated miRNAs (n=36)	mir-184	TGGACGGAGAACTGATAAGGGT	-2.89	0.00E+00	0.00E+00
	mir-15b	CGAATCATTATTTGCTGCTCTA	-2.87	0.00E+00	0.00E+00
	mir-93	ACTGCTGAGCTAGCACTTCCCG	-2.81	6.40E-04	1.86E-03
	let-7a-1//let-7a-3	CTATACAATCTACTGTCTTTC	-2.61	0.00E+00	0.00E+00
	mir-190a	TGATATGTTTGATATATTAGGT	-2.59	0.00E+00	0.00E+00
	mir-33b	GTGCATTGCTGTTGCATTGC	-2.58	0.00E+00	0.00E+00
	mir-6087	TGAGGCGGGGGGGCGAGC	-2.57	1.39E-06	4.98E-06
	mir-18a	TAAGGTGCATCTAGTGCAGATAG	-2.50	0.00E+00	0.00E+00
	mir-592	TTGTGTCAATATGCGATGATGT	-2.49	4.77E-03	1.21E-02
	mir-345	GCTGACTCCTAGTCCAGGGCTC	-2.35	8.24E-07	2.97E-06
	mir-449a	TGGCAGTGTATTGTTAGCTGGT	-2.32	3.10E-14	1.48E-13
	mir-196b	TAGGTAGTTTCCTGTTGTTGGG	-2.32	0.00E+00	0.00E+00
	mir-10b	ACAGATTCGATTCTAGGGGAAT	-2.24	4.84E-04	1.42E-03
	mir-628	TCTAGTAAGAGTGGCAGTCGA	-2.14	2.07E-02	4.55E-02
	mir-24-1//mir-24-2	TGGCTCAGTTCAGCAGGAACAG	-2.11	0.00E+00	0.00E+00
	mir-34a	TGGCAGTGTCTTAGCTGGTTGT	-2.08	0.00E+00	0.00E+00
	mir-374b	CTTAGCAGGTTGTATTATCATT	-2.08	2.72E-04	8.16E-04
	mir-147b	GTGTGCGGAAATGCTTCTGCTA	-2.04	5.46E-07	2.01E-06
	mir-148a	TCAGTGCACTACAGAACTTTGT	-2.03	0.00E+00	0.00E+00
	mir-191	GCTGCGCTTGGATTTCGTCCCC	-2.01	5.73E-07	2.10E-06

6.3.3. MiRNAs Commonly and Exclusively Deregulated in Spheroid-enriched CSCs Models

The differentially expressed miRNAs (p<0.05, FC>2) between the spheroid-enriched CSCs cells derived from MCF-7 and MDA-MB-231 cell were presented using a Venn diagram in order to identify the cluster of miRNAs between different spheroid cell types. As illustrated in Figure 6.5, 102 deregulated miRNAs (15 up and 87 down) were represented in MCF-7 spheroid, while 46 deregulated miRNAs (20 up and 26 down) were exclusively found in MDA-MB-231 spheroid. Most of the deregulated miRNAs observed in spheroid cells (87 miRNAs in MCF-7 spheroid and 26 miRNAs in MDA-MB-231 spheroid) were all down-regulated, suggesting that most of the miRNAs may be tumour suppressors. A total of 20 deregulated miRNAs were found to be commonly regulated between the two cell types, suggesting that these miRNAs may play important roles in maintaining the spheroid-enriched CSCs populations.

The top ten miRNAs that were found to be up-regulated in MCF-7 spheroids revealed that these miRNAs were involved in regulation of cell cycle progression (miR-410) (Lund, 2009), regulation of cellular senescence (miR-127) (Ma, 2010), act as breast cancer biomarkers (miR-411) (Guo, Yuan et al., 2016), modulation of cell apoptosis (miR-1246) (Chien, Domenech et al., 2011), and drug resistance (miR-493) (Chen, Huang et al., 2013).

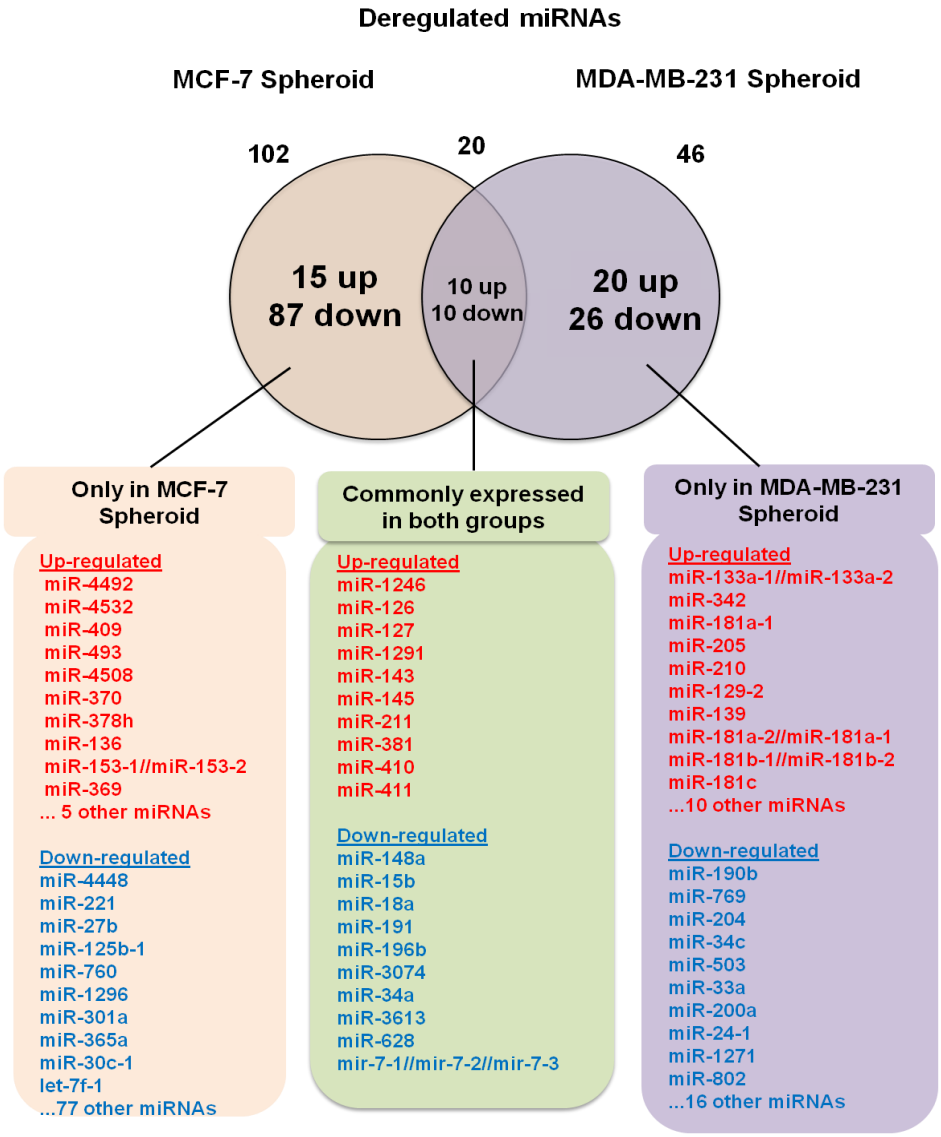


Figure 6.5. Venn diagram showing the number of common and specifically deregulated miRNAs in MCF-7 and MDA-MB-231 spheroids relative to the parental cells. miRNAs differentially expressed between MCF-7 spheroid and parental cells are represented in pink circle, those differentially expressed between MDA-MB-231 spheroid and parental cells are represented in purple circle, and the ones commonly expressed between the two are represented in the overlapped region of the two circles.

6.3.4 Gene-set and Pathway Enrichment Analysis of miRNA Cluster Target Genes

It is well known that miRNAs function by regulating a large number of genes, and each gene is also targeted by many different miRNAs. Therefore, to determine the specific biological functions of the differentially expressed miRNA clusters in the different cell types, target genes, gene annotations and their associated pathways were predicted. Three miRNAs databases available in CyTargetLinker in Cytoscape Version 3.3 including of experimentally validated target genes miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/), TargetScanHuman 6.2 (www.targetscan.org), and microcosm (http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/) were used to predict the potential target genes. The clusters of miRNAs were also subjected for the generation of miRNA-mRNA interaction network. The top 10 significant deregulated miRNAs associated with the uniquely and commonly expressed miRNAs in the spheroid subtypes were visualised as three separate networks using CyTargetLinker in Cytoscape V3.3. These networks provide functionality for representation and integration of mapping targeted genes with the miRNAs visualised by the regulatory interaction network in a graphical manner. Each edge in the network connects a miRNA and its corresponding genes (Figure 6.6 to 6.8). Our results revealed a complex network consisting of the cluster of miRNAs, their targeted genes, enriched pathways and again suggested the complex network of miRNAs-mRNA in the spheroids of different cell types. The annotated GO terms and the enriched pathways in breast CSCs were discussed in Section 6.4.

Gene-set enrichment analysis to elucidate the most relevant Gene Ontology (GO) terms associated with a given gene list of miRNA targets within each miRNA cluster (commonly and exclusively expressed in both spheroids) was performed using DAVID program (http://david.abcc.ncifcrf.gov). The target genes were mapped to the GO annotation dataset for molecular function, biological processes and cellular components. For each GO term, a *p*-value is calculated that represents the probability that observed number of counts could have been resulted from randomly distributing the GO term between the two samples. The resulting list of p-values is then sorted. The GO terms which are most specific for the analysed list of genes will have the lowest p-value. The top 10 statistically significant GO functional categories enriched for each group are represented in Table 6.5, 6.6 and 6.7.

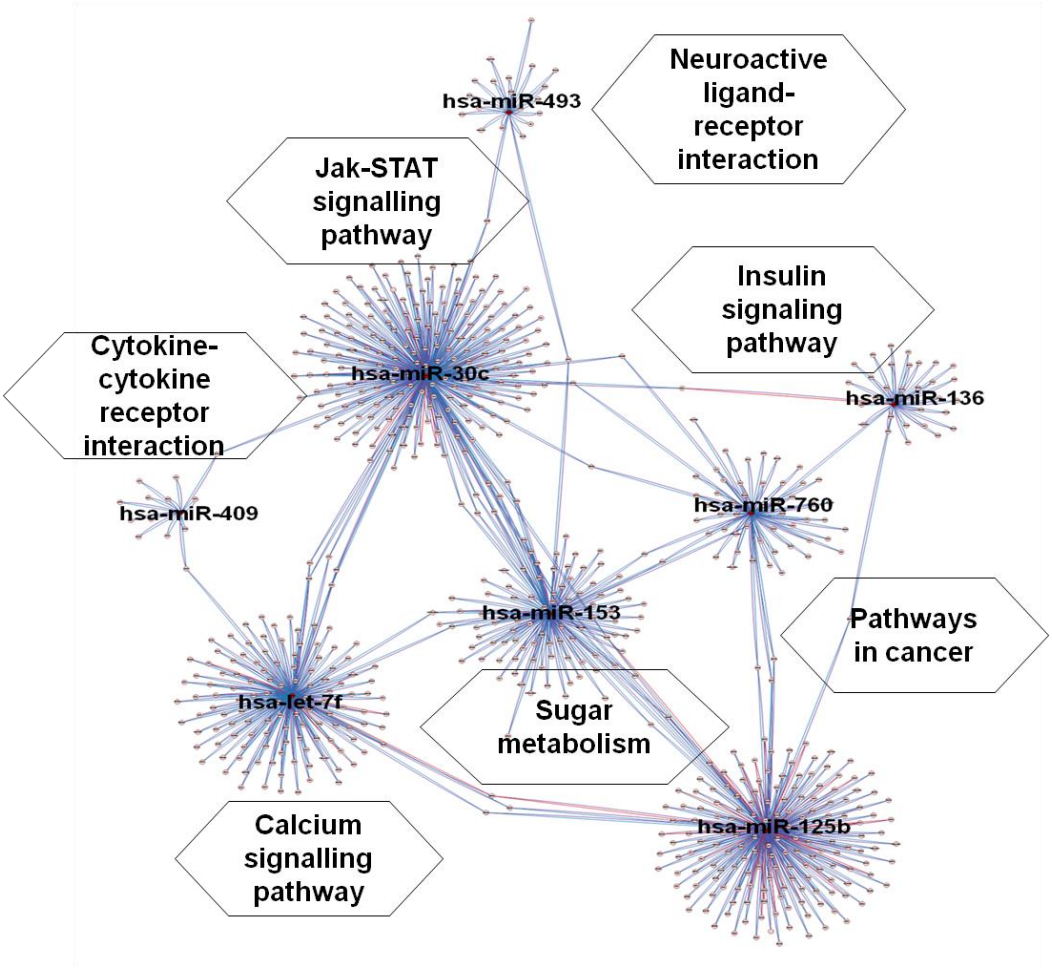


Figure 6.6. Gene networks inferred from differentially expressed miRNAs targets and their important enriched pathways represented in MCF-7 spheroid. Each network includes two types of nodes, red circle represents individual miRNAs and pink circle represents target genes. Targets of the deregulated cluster of miRNAs involved in various pathways based on KEGG ontology.

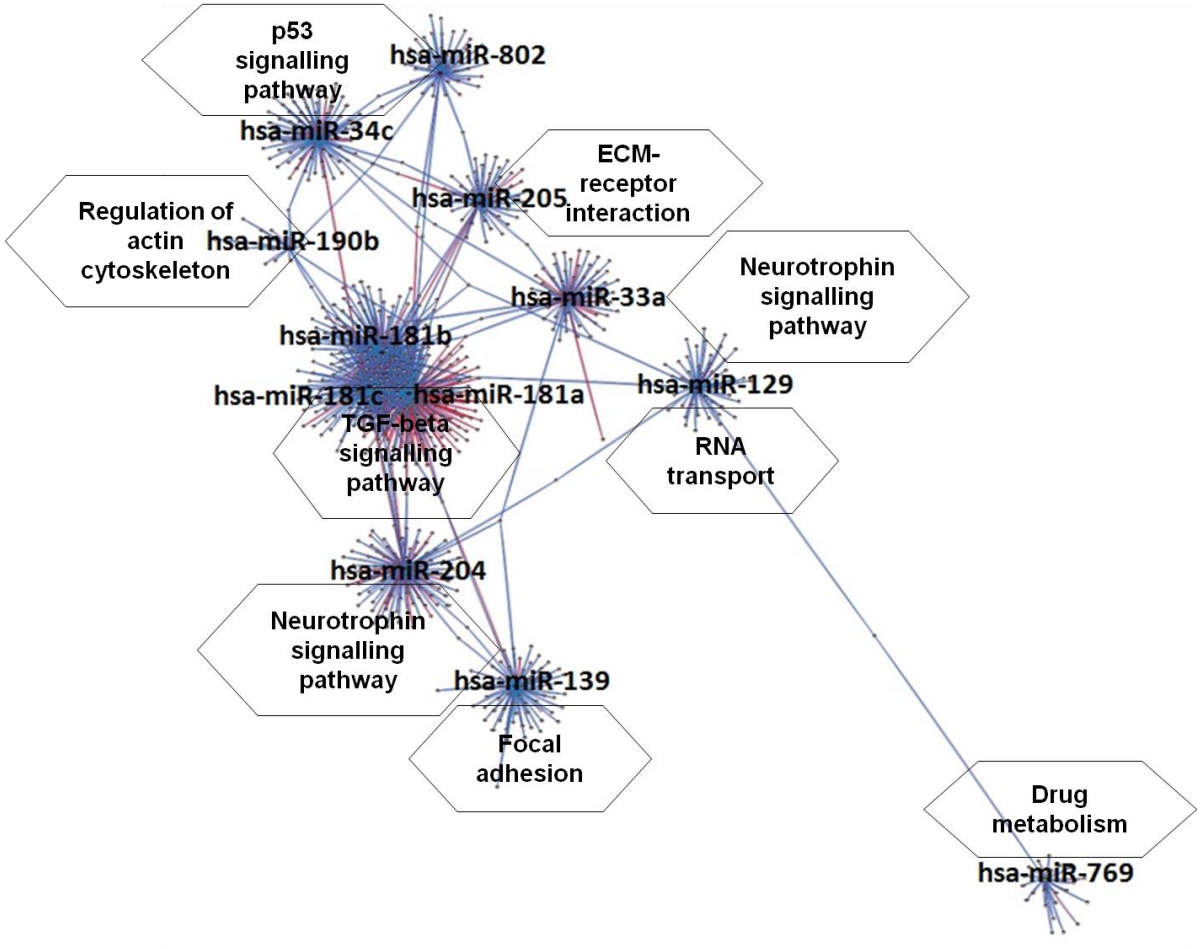


Figure 6.7. Gene networks inferred from differentially expressed miRNAs targets and their important enriched pathways represented in MDA-MB-231 spheroid. Each network includes two types of nodes, red circle represents individual miRNAs and pink circle represents target genes. Targets of the deregulated cluster of miRNAs involved in various pathways based on KEGG ontology.

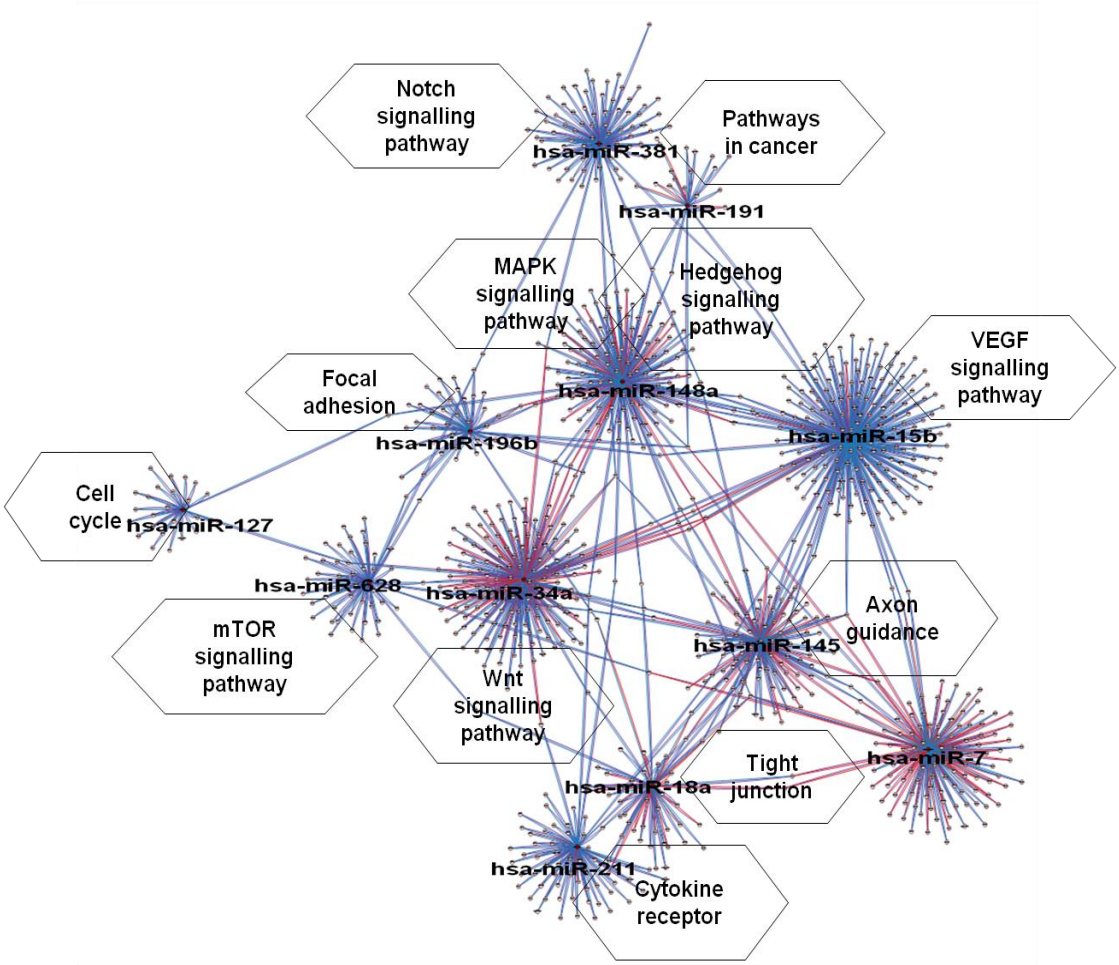


Figure 6.8. Gene networks inferred from differentially expressed miRNAs targets and their important enriched pathways represented in both spheroids. Each network includes two types of nodes, red circle represents individual miRNAs and pink circle represents target genes. Targets of the deregulated cluster of miRNAs involved in various pathways based on KEGG ontology.

The functional analysis revealed that the target genes inferred from the cluster of miRNAs exclusively found in MCF-7 spheroid were significantly associated with a variety of biological processes like regulation of transcription factors, and signal transduction, molecular functions including protein binding and metal ion binding; and the gene products were primarily found in cell cytoplasm, membrane and nucleus (Table 6.5). Analysis of enriched pathways of the target genes for deregulated cluster of miRNAs found in MCF-7 spheroid revealed significant association of JAK-STAT signalling pathway, calcium signalling and insulin signalling pathway (Table 6.8 and Figure 6.9).

Meanwhile, the group of deregulated miRNAs found only in MDA-MB-231 spheroid targeted genes were significantly associated with similar gene annotations and enriched pathways (Table 6.6, Table 6.9 and Figure 6.10). This implied that although different cluster of miRNAs were expressed exclusively in MCF-7 and MDA-MB-231 spheroid, their enriched gene sets and associated pathways were similar. Nevertheless, miRNAs expressed exclusively in MDA-MB-231 spheroid that were associated with neurotrophin signalling pathway, p53 signalling pathway and ECM-receptor interaction were chosen for the validation purpose as these signalling pathways have shown potential for cancer targeted therapy.

A total of 20 miRNAs were found to be commonly deregulated between these two breast cancer spheroid-enriched CSCs cell types, as illustrated in Figure 6.5. The cluster of commonly expressed miRNAs between

the two spheroid cells were further subjected for target predictions purposes in order to obtain more confident target genes for the cluster of miRNAs. Interestingly, the target genes revealed that they were significantly associated with variety of positive regulation processes such as transcription factors, cell proliferation and molecular functions in nucleotide binding, DNA binding and protein kinase binding (Table 6.7). Pathway analysis revealed enrichment in terms related to a set of cancer- and stem cells-related pathways such as Focal adhesion, MAPK, Wnt, Hedgehog, mTOR, and VEGF (Table 6.10 and Figure 6.11). Five miRNAs representing the target genes associated with those pathways were selected for further validation by qPCR expressions analysis.

Table 6.5. Top 10 significant GO terms enriched for each functional gene-set category for predicted miRNAs targets exclusively expressed in MCF-7 spheroid cells. Gene ontology analysis in the biological process (BP), molecular function (MF) and cellular components (CC). Number of genes per annotations indicate enriched levels of genes with modified Fisher exact p-value <0.05. NG: number of genes per annotations.

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Biological Process (BP)	GO:0006355: regulation of transcription, DNA-dependent (BP)	77	1.74E-14	2.97E-11
	GO:0007165: signal transduction (BP)	48	2.60E-07	4.93E-05
	GO:0007275: multicellular organismal development (BP)	44	2.16E-08	7.38E-06
	GO:0055085: transmembrane transport (BP)	41	3.53E-12	3.02E-09
	GO:0045944: positive regulation of transcription from RNA polymerase II promoter (BP)	35	9.45E-10	4.03E-07
	GO:0006811: ion transport (BP)	31	4.29E-08	1.22E-05
	GO:0006915: apoptotic process (BP)	30	7.54E-07	1.07E-04
	GO:0006810: transport (BP)	29	3.11E-06	2.80E-04
	GO:0007268: synaptic transmission (BP)	28	7.59E-10	4.32E-07
	GO:0045893: positive regulation of transcription, DNA-dependent (BP)	27	2.09E-07	4.45E-05

Table 6.5 continued

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Molecular Functions (MF)	GO:0005515: protein binding (MF)	218	7.23E-45	4.35E-42
	GO:0046872: metal ion binding (MF)	118	7.62E-17	1.15E-14
	GO:0000166: nucleotide binding (MF)	103	1.05E-19	3.17E-17
	GO:0005524: ATP binding (MF)	82	6.68E-19	1.34E-16
	GO:0008270: zinc ion binding (MF)	77	1.69E-10	2.04E-08
	GO:0003677: DNA binding (MF)	67	2.57E-08	2.58E-06
	GO:0004872: receptor activity (MF)	59	1.61E-07	1.21E-05
	GO:0003700: sequence-specific DNA binding transcription factor activity (MF)	37	5.08E-06	2.55E-04
	GO:0016787: hydrolase activity (MF)	32	1.16E-03	1.94E-02
	GO:0003676: nucleic acid binding (MF)	26	4.48E-03	4.50E-02

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Cellular components (CC)	GO:0005737: cytoplasm (CC)	221	1.04E-34	3.68E-32
	GO:0005634: nucleus (CC)	219	4.42E-32	7.80E-30
	GO:0016020: membrane (CC)	156	9.06E-20	1.07E-17
	GO:0005886: plasma membrane (CC)	141	8.91E-19	7.86E-17
	GO:0016021: integral to membrane (CC)	141	1.47E-11	1.04E-09
	GO:0005829: cytosol (CC)	83	1.07E-10	6.29E-09
	GO:0005622: intracellular (CC)	64	6.38E-06	1.61E-04
	GO:0005730: nucleolus (CC)	62	9.61E-10	4.24E-08
	GO:0005576: extracellular region (CC)	51	5.00E-03	3.46E-02
	GO:0005887: integral to plasma membrane (CC)	49	1.06E-09	4.14E-08

Table 6.6. Top 10 significant GO terms enriched for each functional gene-set category for predicted miRNAs targets exclusively expressed in MDA-MB-231 spheroid cells. Gene ontology analysis in the biological process (BP), molecular function (MF) and cellular components (CC). Number of genes per annotations indicate enriched levels of genes with modified Fisher exact p-value <0.05. NG: number of genes per annotations.

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Biological Process (BP)	GO:0006355: regulation of transcription, DNA-dependent (BP)	93	2.18E-15	1.68E-12
	GO:0007165: signal transduction (BP)	72	2.63E-13	1.02E-10
	GO:0007275: multicellular organismal development (BP)	60	6.47E-12	1.88E-09
	GO:0045944: positive regulation of transcription from RNA polymerase II promoter (BP)	53	5.56E-17	6.45E-14
	GO:0006915: apoptotic process (BP)	49	5.14E-14	2.99E-11
	GO:0007399: nervous system development (BP)	47	5.22E-19	1.21E-15
	GO:0045893: positive regulation of transcription, DNA-dependent (BP)	42	2.10E-13	9.77E-11
	GO:0030154: cell differentiation (BP)	40	9.35E-11	2.17E-08
	GO:0006810: transport (BP)	38	6.45E-08	9.36E-06
	GO:0055085: transmembrane transport (BP)	37	5.20E-07	4.16E-05

Table 6.6 continued

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Molecular Functions (MF)	GO:0005515: protein binding (MF)	283	1.27E-58	9.51E-56
	GO:0046872: metal ion binding (MF)	161	4.01E-25	1.50E-22
	GO:0000166: nucleotide binding (MF)	127	3.34E-22	8.31E-20
	GO:0008270: zinc ion binding (MF)	110	1.67E-17	2.49E-15
	GO:0005524: ATP binding (MF)	95	1.91E-18	3.57E-16
	GO:0003677: DNA binding (MF)	93	9.43E-13	1.01E-10
	GO:0003700: sequence-specific DNA binding transcription factor activity (MF)	66	5.65E-16	7.04E-14
	GO:0004872: receptor activity (MF)	65	9.44E-06	3.53E-04
	GO:0016787: hydrolase activity (MF)	49	7.30E-07	4.20E-05
	GO:0003676: nucleic acid binding (MF)	41	5.30E-06	2.20E-04

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Cellular components (CC)	GO:0005737: cytoplasm (CC)	296	7.10E-50	2.80E-47
	GO:0005634: nucleus (CC)	281	2.77E-40	5.47E-38
	GO:0016020: membrane (CC)	207	1.87E-27	2.47E-25
	GO:0005829: cytosol (CC)	133	1.30E-24	1.29E-22
	GO:0016021: integral to membrane (CC)	198	8.99E-20	7.11E-18
	GO:0005886: plasma membrane (CC)	169	9.17E-19	6.04E-17
	GO:0005622: intracellular (CC)	95	3.14E-11	1.77E-09
	GO:0005654: nucleoplasm (CC)	56	4.93E-11	2.44E-09
	GO:0005794: Golgi apparatus (CC)	57	2.65E-10	1.16E-08
	GO:0005730: nucleolus (CC)	75	3.50E-10	1.38E-08

Table 6.7. Top 10 significant GO terms enriched for each functional gene-set category for predicted miRNAs targets commonly expressed in both MCF-7 and MDA-MB-231 spheroid cells. Gene ontology analysis in the biological process (BP), molecular function (MF) and cellular components (CC). Number of genes per annotations indicate enriched levels of genes with modified Fisher exact p-value <0.05. NG: number of genes per annotations.

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Biological Process (BP)	GO:0006355: regulation of transcription, DNA-dependent (BP)	104	1.23E-17	1.56E-14
	GO:0007165: signal transduction (BP)	82	6.82E-16	5.79E-13
	GO:0007275: multicellular organismal development (BP)	66	4.98E-13	1.58E-10
	GO:0045944: positive regulation of transcription from RNA polymerase II promoter (BP)	59	4.33E-19	1.10E-15
	GO:0045892: negative regulation of transcription, DNA-dependent (BP)	44	1.09E-15	6.93E-13
	GO:0000122: negative regulation of transcription from RNA polymerase II promoter (BP)	42	8.65E-14	3.67E-11
	GO:0008284: positive regulation of cell proliferation (BP)	41	3.17E-15	1.62E-12
	GO:0045893: positive regulation of transcription, DNA-dependent (BP)	41	1.66E-11	3.53E-09
	GO:0006915: apoptotic process (BP)	40	5.63E-08	5.51E-06
	GO:0006468: protein phosphorylation (BP)	38	7.24E-12	1.84E-09

Table 6.7 continued

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Molecular Functions (MF)	GO:0005515: protein binding (MF)	323	2.82E-71	2.14E-68
	GO:0046872: metal ion binding (MF)	155	2.53E-18	3.19E-16
	GO:0000166: nucleotide binding (MF)	134	1.18E-21	2.98E-19
	GO:0003677: DNA binding (MF)	117	3.82E-20	5.78E-18
	GO:0005524: ATP binding (MF)	107	2.18E-21	4.12E-19
	GO:0008270: zinc ion binding (MF)	106	4.90E-13	4.64E-11
	GO:0003700: sequence-specific DNA binding transcription factor activity (MF)	84	4.09E-24	1.55E-21
	GO:0004872: receptor activity (MF)	60	2.02E-03	2.06E-02
	GO:0016787: hydrolase activity (MF)	58	5.65E-09	3.29E-07
	GO:0043565: sequence-specific DNA binding (MF)	53	1.91E-16	2.07E-14

	Annotation Term	NG	Unadjusted p-value	FDR p-value
Cellular components (CC)	GO:0005737: cytoplasm (CC)	349	7.95E-68	3.15E-65
	GO:0005634: nucleus (CC)	340	5.37E-60	1.06E-57
	GO:0016020: membrane (CC)	204	5.86E-21	4.65E-19
	GO:0005886: plasma membrane (CC)	192	5.89E-23	5.85E-21
	GO:0016021: integral to membrane (CC)	182	8.65E-11	3.43E-09
	GO:0005829: cytosol (CC)	147	1.98E-27	2.61E-25
	GO:0005622: intracellular (CC)	100	1.02E-10	3.69E-09
	GO:0005730: nucleolus (CC)	87	6.80E-13	3.86E-11
	GO:0005654: nucleoplasm (CC)	74	1.76E-18	1.16E-16
	GO:0005783: endoplasmic reticulum (CC)	67	2.16E-12	9.55E-11

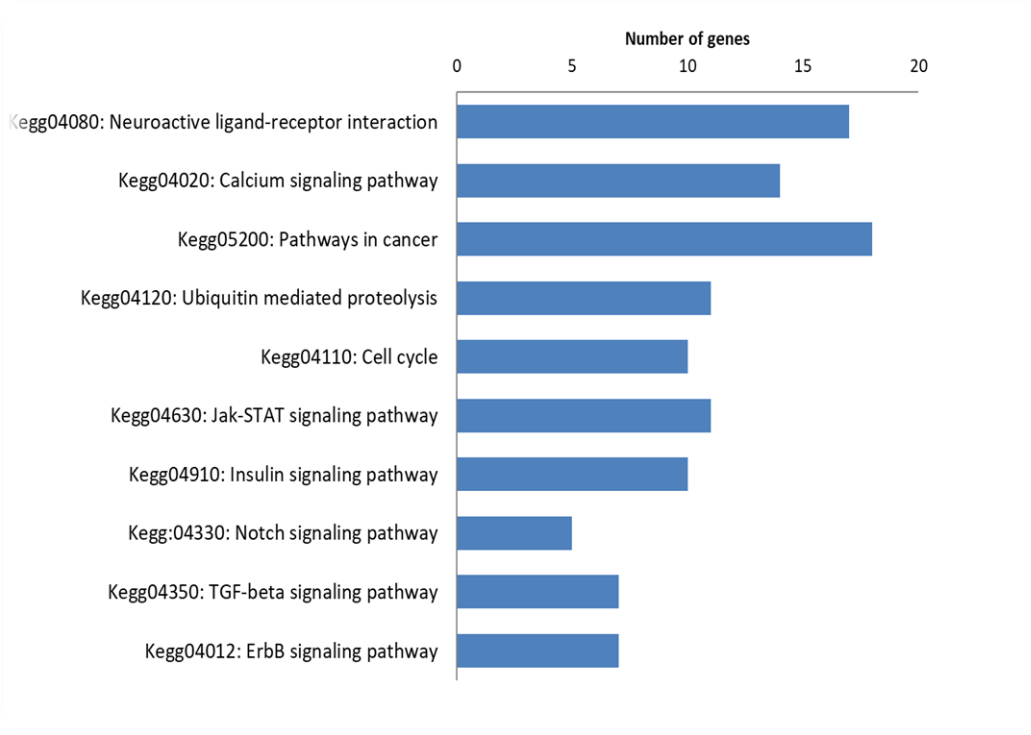


Figure 6.9. Significantly enriched KEGG pathways associated to the gene targets of the exclusively differentially expressed miRNAs in MCF-7 spheroid cells using DAVID program. KEGG, Kyoto Encyclopaedias of Genes and Genomes.

Table 6.8. Significantly enriched pathways associated to the gene targets of the exclusively differentially expressed miRNAs in MCF-7 spheroid cells using DAVID program. The table shows the list of enriched KEGG pathways with the corresponding corrected p-value based on hypergeometric test with the most significant pathways listed from top down.

Pathways	***p*-value**
Kegg04080: Neuroactive ligand-receptor interaction	2.96E-04
Kegg04020: Calcium signalling pathway	3.88E-04
Kegg05200: Pathways in cancer	6.50E-04
Kegg04120: Ubiquitin mediated proteolysis	6.74E-04
Kegg04110: Cell cycle	1.17E-03
Kegg04630: Jak-STAT signalling pathway	1.49E-03
Kegg04910: Insulin signalling pathway	1.94E-03
Kegg04350: TGF-beta signalling pathway	5.77E-03
Kegg04012: ErbB signalling pathway	7.85E-03

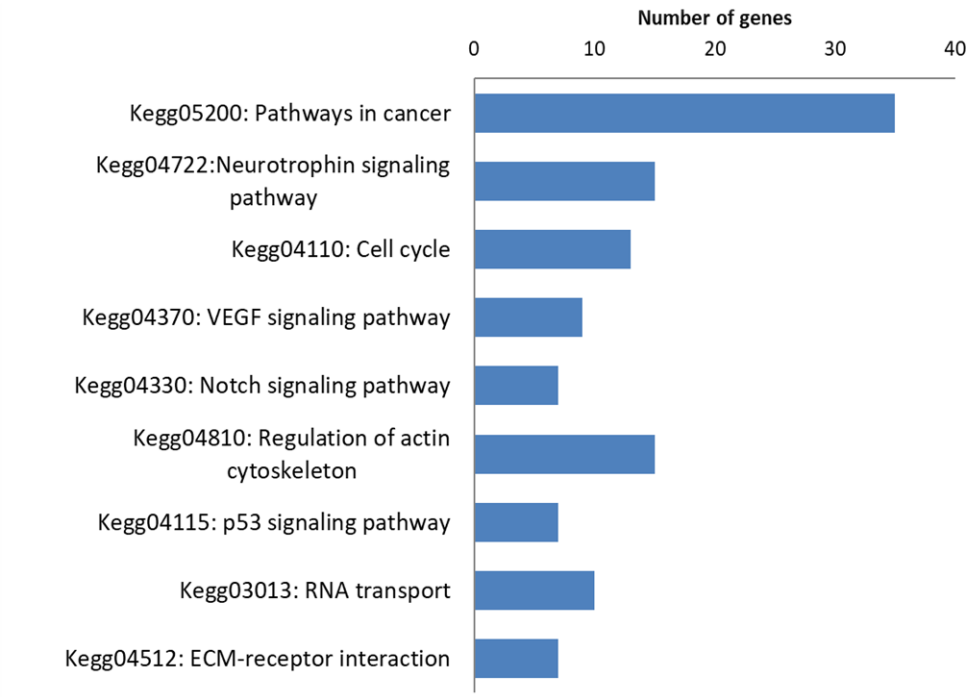


Figure 6.10. Significantly enriched KEGG pathways associated to the gene targets of the exclusively differentially expressed miRNAs in MDA-MB-231 spheroid cells using DAVID program. KEGG, Kyoto Encyclopaedias of Genes and Genomes.

Table 6.9. Significantly enriched pathways associated to the gene targets of the exclusively differentially expressed miRNAs in MDA-MB-231 spheroid cells using DAVID program. The table shows the list of enriched KEGG pathways with the corresponding corrected p-value based on hypergeometric test with the most significant pathways listed from top down.

Pathways	***p*-value**
Kegg05200: Pathways in cancer	1.87E-11
Kegg04722: Neurotrophin signalling pathway	4.82E-06
Kegg04110: Cell cycle	6.50E-05
Kegg04370: VEGF signalling pathway	0.00050687
Kegg04330: Notch signalling pathway	0.000656646
Kegg04810: Regulation of actin cytoskeleton	0.000948353
Kegg04115: p53 signalling pathway	0.00446784
Kegg03013: RNA transport	0.00845516
Kegg04512: ECM-receptor interaction	0.0113544

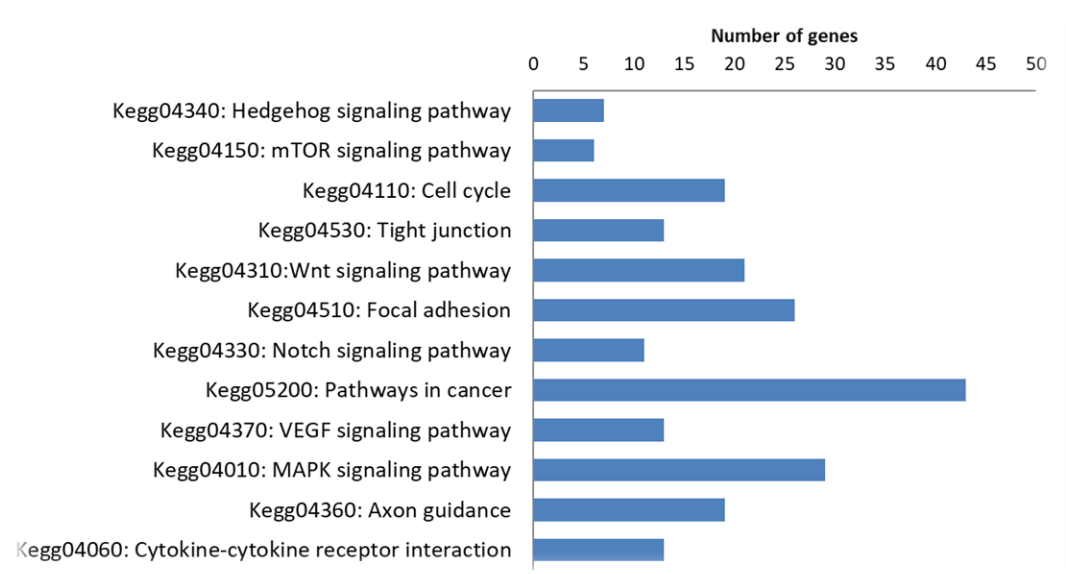


Figure 6.11. Significantly enriched KEGG pathways associated to the gene targets of the commonly differentially expressed miRNAs in both MCF-7 and MDA-MB-231 spheroid cells using DAVID program. KEGG, Kyoto Encyclopaedias of Genes and Genomes.

Table 6.10. Significantly enriched pathways associated to the gene targets of the commonly differentially expressed miRNAs in both MCF-7 and MDA-MB-231 spheroid cells using DAVID program. The table shows the list of enriched KEGG pathways with the corresponding corrected p-value based on hypergeometric test with the most significant pathways listed from top down.

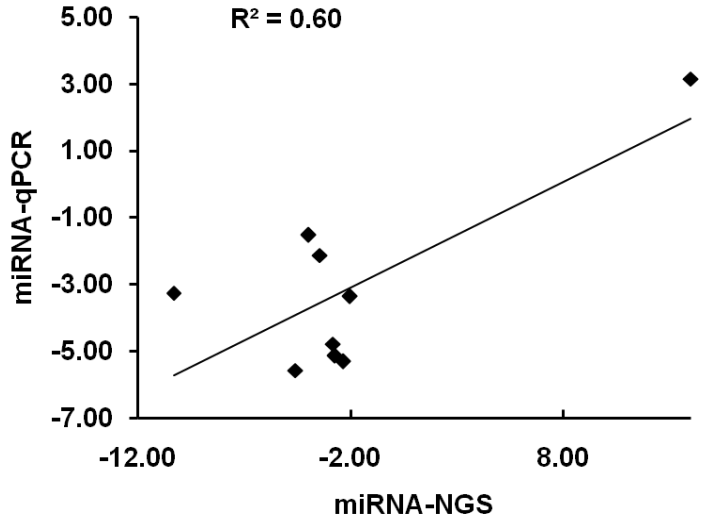
Pathways	*p*-value
Kegg04340: Hedgehog signaling pathway	1.18E-08
Kegg04150: mTOR signaling pathway	1.94E-06
Kegg04110: Cell cycle	7.53E-06
Kegg04530: Tight junction	1.44E-05
Kegg04310: Wnt signaling pathway	7.01E-05
Kegg04510: Focal adhesion	2.82E-04
Kegg04330: Notch signaling pathway	5.04E-04
Kegg05200: Pathways in cancer	4.02E-03
Kegg04370: VEGF signaling pathway	7.64E-03
Kegg04010: MAPK signaling pathway	9.46E-03
Kegg04360: Axon guidance	1.64E-02
Kegg04060: Cytokine-cytokine receptor interaction	3.74E-02

6.3.5 Data Validation of The Differentially Expressed miRNAs Using Real-time PCR

In order to validate the expression changes of the selected miRNAs obtained using NGS, a different quantitative approach is required. Therefore, real-time PCR which is the gold standard method was used to validate the miRNA-NGS results. A total of twelve dysregulated miRNAs that exhibited a more than 2-fold change in expression were selected for validation purposes. The primer sequences for real-time PCR are listed in Section 3.8. The miRNAs were selected based on statistical significance ($p<0.05$) and their important roles in biological or signalling pathways altered in breast cancer spheroid cells with respect to their controls (parental cells) culture condition. Fold changes of the real-time PCR were obtained from the mean expression value of the Cq (quantitation cycles) values normalised to the reference miRNAs. The relationship of the data between the NGS and real-time PCR was also determined using the method. Correlation plot is used to measure the degree of variation between one set of data to another (Git, Dvinge et al., 2009, Lee, Hong et al., 2014). The log2 ratios of the miRNAs obtained using NGS were plotted in accordance to the corresponding real-time values (Figure 6.12). The measurement of the correlation coefficient R squared, R^2, was determined by linear regression analysis. The measurements of the miRNA expression in the both cell types determined by the NGS and real-time methods showed good correlation ($R^2> 0.60$) with a better correlation showed in MDA-MB-231 group (Figure 6.12B). As illustrated in Figure 6.13 and Appendix L, the levels of the selected miRNAs measured by real-time PCR and NGS showed similar trend, indicating the reliability of the sequencing data and supporting the

interpretation of the expression profiles and pathways information based on the miRNAs expression results in this study.

A



B

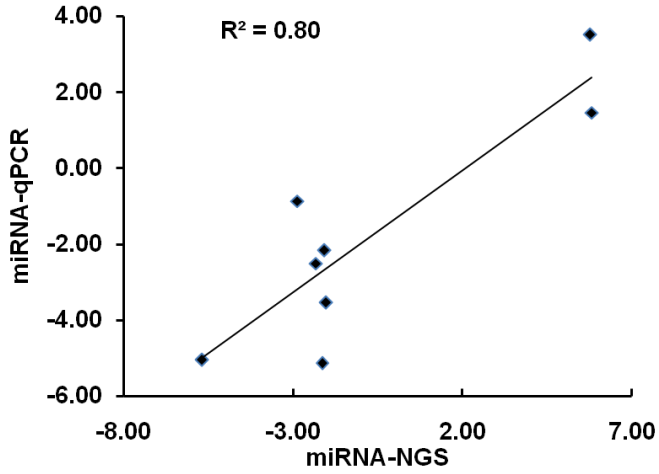


Figure 6.12. Correlation of real-time PCR and NGS data. Scatter plots shows the correlation of the ratios from the NGS and real-time PCR data sets. miRNAs that were differentially expressed between the NGS analysis were selected and validated with the same samples by real-time PCR analysis. Ratios of expression from both datasets correlated well with R^2 >0.6 determined by linear regression analysis. (A) MCF-7 cell type; (B) MDA-MB-231 cell type. R^2: correlation coefficient.

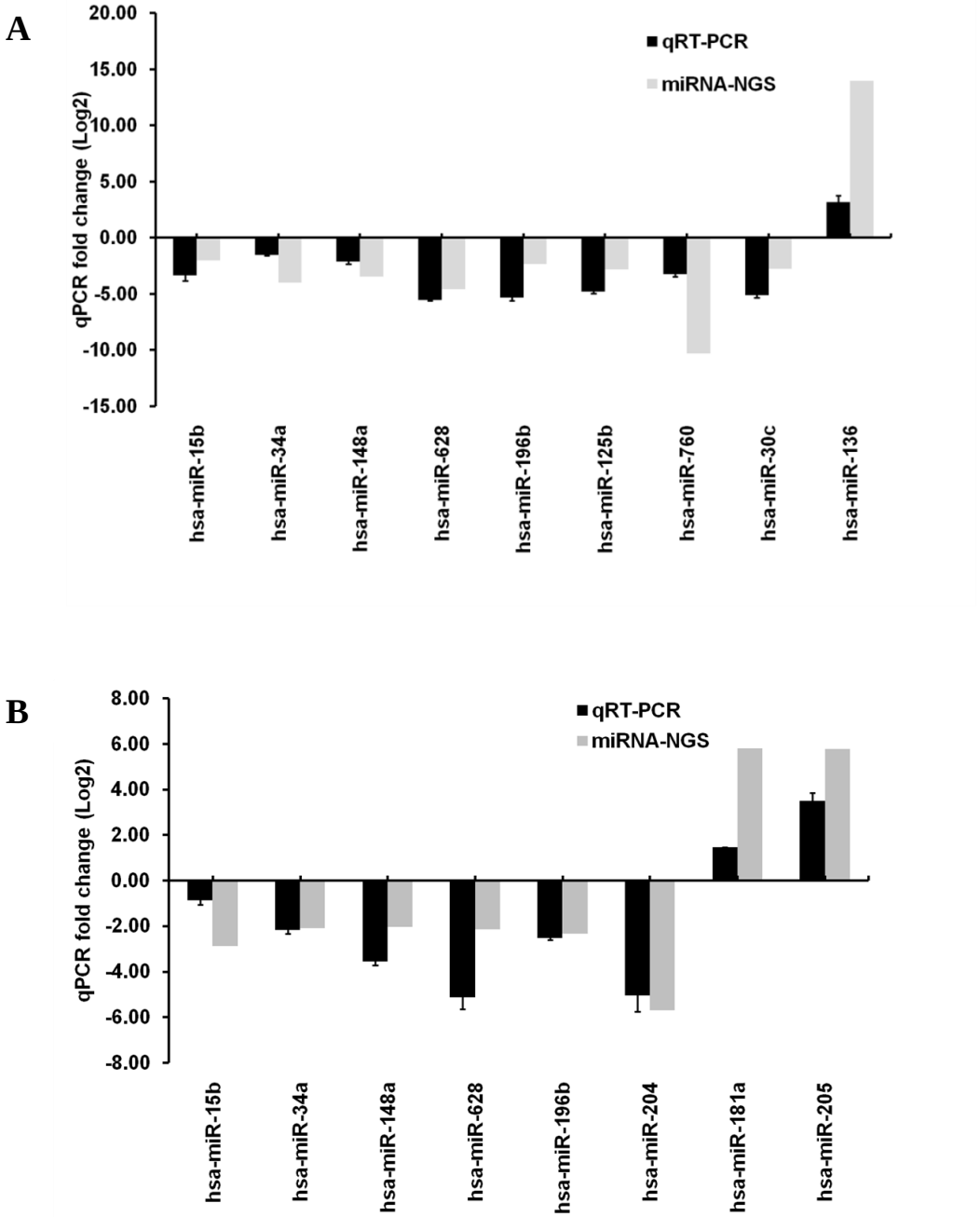


Figure 6.13. Validation of known miRNAs with real-time PCR. Comparison of the real-time PCR and miRNA-NGS sequencing data log2-fold change for the selected miRNAs between the spheroids and parental cells of both breast cancer cell lines. A similar expression trends from both NGS and real-time PCR were observed. (A) qPCR validation for MCF-7 cell type; (B) qPCR validation for MDA-MB-231 cell type.

6.4 Discussion

6.4.1 Up-regulated miRNAs in MCF-7 Spheroid-enriched CSCs

Herein, the Venn diagram showed 102 deregulated miRNAs exclusively expressed in MCF-7 spheroids (as compared to control MCF-7 parental monolayer cells); of these 15 were up-regulated and 87 were down-regulated.

Of the top ten miRNAs that were found to be up-regulated, seven miRNAs that have been previously been reported to be involved in breast cancer were identified; miR-493, miR-4508, miR-370, miR-378h, miR-136, miR-153, and miR-369 (Table 6.11). Literature analysis have demonstrated that these seven miRNAs are closely associated to the invasion and metastasis of breast cancer through a variety of target genes and signalling pathways.

Increased miR-493 has also been explored in cancers in previous studies, but not in CSCs. Upregulation of miR-493 was found to inhibit proliferation and metastasis of gastric cancer cells (Zhou, W., et al., 2015). Also, it was also demonstrated that overexpression of miR-493 prevented liver metastasis through induction of cell death of metastasized cells (Okamoto, K., et al., 2012). Though, miR-493 has not been previously reported in CSCs, the high expression of this miRNA found in this study could be used as potential therapeutic tool to improve clinical outcomes in breast cancer. Studies have shown that elevated expression of miR-493 inhibited the protein expression of FUT4, which in turn decrease the breast cancer invasion and tumorigenicity.

The underlying mechanism of FUT4 pathway regulated by miR-493 could serve as a therapeutic potential in breast cancer (Zhao, Feng et al., 2016).

Two miRNAs, miR-4508 and miR-370, were previously reported in breast cancer. For instance, miR-4508 is up-regulated in BT474/HER2+ breast cancer cell line, and down-regulated in MCF-7/HER2- breast cancer cell line (X Wang, D Jiang et al., 2017). Phenotypically, BT474 is a luminal B cell type which in principle is more invasive and consequently more aggressive than luminal A cells, such as MCF-7 (Holliday and Speirs, 2011). Consistent with our study, upregulated miR-4508 in MCF-7 spheroid cells suggested the transition of the weakly luminal A cells into more aggressive basal features. Thus, this could explain the enhanced migration and invasion abilities found in MCF-7 spheroid relative to their parental cells. Therefore, miR-4508 could be an important biomarker to capture the EMT process of luminal subtype to luminal-like CSCs cells. Furthermore, miR-370 playing a pivotal role as an oncogenic miRNA was found to be upregulated in breast cancer malignant tumours (Mollainezhad, Eskandari et al., 2016). The study indicated that overexpression of miR-370 correlated with ER and PR status, suggesting its significance use as prognostic biomarker in advanced breast cancer.

Notably, aberrant expressions of the panel of exclusively deregulated miRNAs in MCF-7 spheroid were also previously reported in other cancers as well such as pancreatic, colorectal, prostate and lung cancer and play important roles in tumorigenesis of these tumours (Hao, Ding et al., 2017, Ho, Noor et al., 2018, Lo, Hung et al., 2012, Ren, Qi et al., 2017, Wang and Liu, 2016,

Zhang, Luo et al., 2014, Zhi, Zhao et al., 2017). This is similar to the results in our study as these miRNAs were up-regulated in MCF-7 spheroid enriched CSCs models compared to their parental counterparts that could possibly promote the generation of breast epithelial-like CSCs. However, there were conflicting results as several miRNAs which were found to be up-regulated in this study were down-regulated in other studies (Cui, Zhang et al., 2013, Gu, Cheng et al., 2014). For instance, down regulation of miR-378h was found to suppress cell proliferation and induce apoptosis in colorectal cancer cells (Wang, Z., et al., 2015) and also observed in prostate cancer associated with aggressive disease phenotype (Avgeris, M., 2014). Thus, it suggests that miR-378h targets might have different role in progression or inhibition of breast CSCs. Therefore, the contradictory effect of miR-378h on various cancers may due to the ability of miRNA to regulate the expression of multiple target genes that have different function.

Two miRNAs (miR-4492 and miR-4532), which have not been previously reported in breast cancer had a FC >10. Judging from such high degree of differences, these miRNAs are worthy candidates to further explore novel biological links towards this subtype of breast cancer research.

6.4.2 Down-regulated miRNAs in MCF-7 Spheroid-enriched CSCs

The five down-regulated miRNAs found in breast cancer have roles in drug resistance and EMT (miR-221) (Xu, Ohms et al., 2013), tumour growth (miR-27b) (Hoppe, Achinger-Kawecka et al., 2013), cell motility (miR-125b)

(Josson, Gururajan et al., 2014), chemoresistance (miR-760) (Zhang, Luo et al., 2014), and cell migration, invasion and proliferation (miR-301a) (Gan, Yang et al., 2014). Hence, this provides a reason the MCF-7 spheroids behave more aggressively and exhibit higher chemoresistance compared to the parental. MiR-221, showing a >50-fold reduction in expression, is particularly interesting. It has been reviewed to be involved in oestrogen receptor-negative breast cancer and also involved in drug resistance (Akhavantabasi, Sapmaz et al., 2012). MiR-221 has been identified as a regulator of EMT transition and induced EMT-like characteristics, responsible for cell invasion and migration (Lv, Xia et al., 2014), which is consistent with the biological significance of the present study.

MiR-760 which had been functionally explored as a biomarker for early detection of colorectal cancer (Shi, Gerster et al., 2011) and breast cancer (Zhang, Man et al., 2014), was also found as one of the top down-regulated miRNAs in this study. These findings indicated that miR-221 and miR-760 may be associated with luminal-like spheroid cancer and could be possibly be an important CSC-regulated miRNA which have not been previously reported. Down-regulation of let-7f and miR-30c confirmed the notion that CSCs were enriched in MCF-7 spheroids (Schwarzenbacher, Balic et al., 2013, Shah and Calin, 2011, Stinson, Lackner et al., 2011) and these miRNAs may not only regulate CSCs mechanisms, but may also enhance breast cancer cells sensitivity to chemotherapeutics drugs. MiR-125b-5p was favoured to have tumour suppressor roles by targeting ENPEP gene, where enforced expression in MCF-7 cells reduced cell proliferation and anchorage-independent growth

by positive regulation in breast tumorigenesis (Feliciano, Castellvi et al., 2013). Downregulated expression of miR-125b-5p was also found to increase HER2 protein expression, leading to worse prognostic outcome in luminal A breast cancer patients (Bailey, Westerling et al., 2015). Moreover, deregulated expression of miR-760 was previously reported as one of the miRNAs regulated by oestrogen-responsive gene clusters, suggesting its potential use as biomarkers for the luminal-like breast cancer cells (Cicatiello, Mutarelli et al., 2009).

In another similar study, expression of miR-30c has been linked to chemotherapy resistance specifically in luminal A breast tumours by regulating TWF1 and IL-11 genes in the miR-30c-mediated pathway (Bockhorn, Dalton et al., 2013) and deregulation of miR-30c has led to increase in cell proliferation, drug sensitivity and cancer progression. As for miR-136-5p, although it has not been to be associated with luminal-like spheroid breast cancer, it was previously reported to be deregulated in TNBC responsible for increased tumour invasion and metastasis (Yan, Li et al., 2016). Furthermore, deregulation of miR-136 has also been shown in other cancers, including lung cancer (Shen, Yue et al., 2013) and ovarian cancer (Zhao, Liu et al., 2015), suggesting the potential use as biomarker in cancers.

Two miRNAs, miR-4448 and miR-1296, which have not been linked to breast cancer, but in other types of cancer were observed in our study (Biankin, Waddell et al., 2012, Yu, Yao et al., 2007). Further studies are needed to

elucidate these potential novel oncogenic miRNAs implicated in luminal-like breast cancer and their functional roles in modulating CSCs.

6.4.3 Gene Ontology Analysis of the Uniquely Differentially Expressed miRNAs in MCF-7 Spheroid-enriched CSCs

GO analysis also showed noticeable changes in the cellular component of the MCF-7 spheroid-enriched CSCs models including ‘integral to membrane’, ‘intracellular’ and ‘membrane’ which gives the spheroids the 3D structure and which also could contribute to their higher cell proliferation and migration characteristics. Gene clustering governing cell cycle is also an important feature of CSCs for the maintenance of self-renew ability and proliferation characteristics (White and Dalton, 2005). GO analysis demonstrated that there are changes in cell cycle-related gene cluster including, ‘regulation of transcription, DNA dependent’, ‘signal transduction’, and ‘positive regulation of transcription from RNA polymerase II promoter’ which correlates to their higher proliferation activity and metastasis ability in spheroids.

Taken together, the GO terms enriched from the predicted miRNAs found in the spheroids of luminal subtype further illustrated the roles of these miRNAs in breast spheroid chemoresistance and maintenance of self-renewal. GO enrichment in molecular function and cellular component showed more changes in MCF-7 spheroid cells than parental cells, supporting the biological

significance of the results that spheroid-enriched CSCs cells behave differently during the EMT transition stage. It appeared likely that the spheroids have undergone some morphological changes and through the EMT process, proliferated faster, acquired a higher migration and invasion properties, higher chemosensitivity and most likely enriched with CSCs populations, in agreement with other recent studies (White and Dalton, 2005, Yim and Sheetz, 2012).

Table 6.11. The expression profiles of all the top ten miRNAs found in this study and matched to the existing literatures. Uniquely up- and down-regulated differentially expressed miRNAs in MCF-7 spheroids when compared to their parental counterparts

	miRNA	Deregulation	Tissue type	Validated target (s)	Mechanism	References
Uniquely expressed in MCF-7 spheroids in relative to parental cells	**Breast CSCs-related miRNAs found in this study**					
	miR-30c	Downregulated	Breast	NOV/CCN3	Invasion and metastasis	(Dobson, Taipaleenmäki et al., 2014)
	let-7f	Downregulated	Breast	CCR7	Migration and invasion	(Kim, Shin et al., 2012)
	miR-221	Downregulated	Breast	PTEN	EMT	(Li, Lu et al., 2016)
	Breast cancer-related miRNAs found in this study					
	miR-493	Upregulated	Breast	FUT4	Invasion and metastasis, biomarker	(Gasparini, Cascione et al., 2014, Zhao, Feng et al., 2016)
	miR-4508	Upregulated	Breast	HER-2	Biomarker	(X Wang, D Jiang et al., 2017)
	miR-370	Upregulated	Breast	N/A	Biomarker	(Mollainezhad, Eskandari et al., 2016)
	miR-378h	Upregulated	Breast	PGC-1β	Proliferation	(Eichner, Perry et al., 2010)
	miR-136	Upregulated	Breast	RASAL2	Invasion and metastasis	(Yan, Li et al., 2016)
	miR-153	Upregulated	Breast	HECTD3	Apoptosis	(Wu, Li et al., 2016)
	miR-369	Upregulated	Breast	N/A	EMT	(Hao, Zhang et al., 2014)
	miR-27b	Downregulated	Breast	CBLB	Drug resistance and biomarker	(Chen, Si et al., 2018)
	miR-125b	Downregulated	Breast	ENPEP ARID3B	Tumorigenesis Migration and motility	(Feliciano, Castellvi et al., 2013) (Akhavantabasi, Sapmaz et al., 2012)
	miR-760	Downregulated	Breast	N/A	Drug resistance	(Hu S.-H. , Wang C.-H. et

					al., 2016)
	miR-301a	Downregulated	Breast	N/A	Biomarker for poor prognosis	(Zheng, Huang et al., 2018)
	Other cancers-related miRNAs found in this study				
	miR-4492	Upregulated	Brain	N/A	Cell adhesion	(Ahn, Ahn et al., 2018)
	miR-4532	Upregulated	Oesophageal	N/A	Biomarker	(Drahos, Schwameis et al., 2016)
	miR-409	Upregulated	Prostate	N/A	Tumour initiation and EMT	(Josson, Gururajan et al., 2014)
	miR-4448	Downregulated	Gastric	EZH2	Migration	(Hibino, Saito et al., 2014)
	miR-1296	Downregulated	Gastric	ERBB2	Migration and invasion	(Shan, Wen et al., 2017)
	miR-365a	Downregulated	Lung	FOXC2	Invasion and metastasis	(Wang, Zheng et al., 2018)

N/A: Not available

6.4.4 Up-regulated miRNAs in MDA-MB-231 Spheroid-enriched CSCs

Our miRNA transcriptomic profiling suggests that both of the spheroid-enriched CSCs derived from both cell lines may regulate different group of miRNAs. The differences may also be attributed by the status of their EMT, metastasis and their chemoresistance properties. Herein, the Venn diagram showed 46 deregulated miRNAs exclusively expressed in MDA-MB-231 spheroids (as compared to control MDA-MB-231 parental monolayer cells); of these 20 were up-regulated and 26 were down-regulated.

All the top ten overexpressed miRNAs observed in MDA-MB-231 spheroids have been previously reported in breast cancer as well as some other cancers such as pancreatic, gastric, liver, colorectal and lung. Some of the miRNAs uniquely found in MDA-MB-231 spheroids in this study reflect its myoepithelial origins. For example, miR-205-5p preferentially expressed in myoepithelial cells is normally found in triple negative breast tumours. Deregulation of miR-205-5p targets E2F1, which is a key player in cell cycle progression, promotes cell proliferation, cell progression and metastasis; this is in agreement with other studies (Dar, Majid et al., 2011, Wu, Zhu et al., 2009). Besides cell cycle pathway, the potential pathway of miR-205 included EM-receptor interaction pathway which targets LAMC1, the main component in cellular matrix and cell adhesion, and their deregulation coupled with defective p53 status enhances cancer metastasis (Piovan, Palmieri et al., 2012). In addition, miR-205 preferentially expressed in basal-like breast cancer cells is normally found in triple negative breast tumours and a key player in cell cycle

progression where it regulates cell proliferation, cell progression and enhance metastasis (Elgamal, Park et al., 2013, Radojicic, Zaravinos et al., 2011).

Particularly interesting miRNA is the miR-181 family, which consists of miR-181a, -b, and -c. MiR-181a which was upregulated in the MDA-MB-231 spheroids, is also commonly found to be overexpressed in TNBC tissues (Taylor, Sossey-Alaoui et al., 2012). High expression of miR-181a has been known to inhibit TGFβR3 protein translation which subsequently increases metastasis, invasion, migration, and reverting anoikis resistance in breast cancer cells through their negative regulation in TGF-beta signalling pathway.

Another particularly interesting miRNA is the miR-342, which is involved in oestrogen receptor sensitivity was found to be up-regulated in this type of spheroids, and commonly associated with oestrogen receptor expression of breast cancer cells and tissues (He, Wu et al., 2013). Up-regulated expression of miR-342 sensitised the MCF-7 cells to tamoxifen-induced apoptosis and reduced cell proliferation. However, miR-342 was found to increase MDA-MB-231 spheroid chemoresistance and increase cellular proliferation, suggesting miR-342 regulation on ER receptors is not completely understood. Similarly, overexpression of miR-342 has been reported in MDA-MB-231 cell line, by inhibiting ID4 gene and increase of BRCA1 expression through negative regulation, contributing to malignancy promotion and maintenance of cancer stem cells (Crippa, Lusa et al., 2014).

On the other hand, the sole up-regulated miRNA associated with solid tumour was miR-210 that was implicated as the master of hypoxia-response regulator (Chan, Banerjee et al., 2012). Closer examination revealed the functional link between hypoxia induction in tumour microenvironment and the up-regulation of miR-210, which could perhaps perturb their intrinsic regulation associated with metastatic stage of breast cancer (Grosso, Bella et al., 2017, Qin, Furong et al., 2014).

6.4.5 Down-regulated miRNAs in MDA-MB-231 Spheroid-enriched CSCs

In contrast, 26 miRNAs were found to be significantly down-regulated in MDA-MB-231 spheroid-enriched CSCs model. Of the top ten down-regulated miRNAs, five miRNAs (miR-769, miR-204, miR-34c, miR-33a, and miR-200a,) have been previously reported to be dysregulated in breast cancer tissues or cells.

These miRNAs have roles in promoting cell proliferation and reduce apoptosis (miR-769, and miR-34c), forming colony (miR-204), increasing tumour invasion and migration (miR-769 and miR-204) and maintaining stemness (miR-33a and miR-200a) either through positive or negative regulation of their target genes in several studies as shown in Table 6.12. Notably, miR-200a, which was 3-fold down regulated in this study, had been widely demonstrated as a powerful regulator of epithelial-mesenchymal transition (EMT) (Park, Gaur et al., 2008, Shimono, Zabala et al., 2009). Increasing evidences have suggested that cells with EMT traits share many

biological characteristics with CSCs. Inhibition of miR-200a promote expression of ZEB1 and ZEB2, and subsequently induced EMT and metastasis (Lu, Lu et al., 2014).

MiR-204 which has been reported as down-regulated in a few cancers (Yin Y, Zhang B et al., 2014), was also present at low levels in theMDA-MB-231 spheroid and their reduced expression of miR-204 was associated with poor clinical outcome in basal-like TNBC, suggesting their potential diagnostic use (Li, Jin et al., 2014). In contrast, ectopic expression of miR-204 has been demonstrated to restore anoikis sensitivity and reduced the invasiveness and metastatic behaviour in ovarian cancer cell line (Yan H, Wu W et al., 2015). Therefore, targeting miR-204 could possibly enhance anoikis sensitivity of basal-like breast cancer via neutrotrophin signalling pathway, thus making it a possible target therapy for metastasis breast cancer.

Similarly, miR-33a, which was down-regulated, was found to maintain stemness by up-regulation of the target genes such as HMGA2, SALL4 and TWIST1 promoting the self-renewal and also the migration and invasion in breast cancer (Lin, Liu et al., 2015). Additionally, the expression level of miR-24-1 in MDA-MB-231 spheroids was 3 times lower than in the parental cells in the present study. A recent study has shown that the deregulation of miR-24-1 promoted spheroid formation via targeting BimL, an apoptotic gene, and induced chemoresistance, highlighting the importance of this miRNA as new therapeutic target for basal cell-type specific miRNA (Roscigno, Puoti et al., 2017). MiR-190b has been reported to mediate EMT by activating ZEB1

through EMT-inducing TGF-ß signalling pathway (Yu, Luo et al., 2018). Similarly, the present study showed downregulation of miR-190b in MDA-MB-231 spheroid derived CSCs, which suggesting EMT activation in the cells.

6.4.6 Gene Ontology Analysis of The Uniquely Differentially Expressed miRNAs in MDA-MB-231 Spheroid-enriched CSCs

Functional annotations through GO analysis demonstrated that the gene sets inferred from the cluster of miRNAs exclusively expressed in MDA-MB-231 spheroids were enriched in biological processes pathways such as 'transport' and 'transmembrane transport'; molecular functions including 'nucleotide binding' and 'protein binding'; and as well as cellular characteristics such as 'Golgi apparatus' and 'integral to membrane'. These predicted functional annotations are consistent with previous literatures reported that they contribute to the higher metastasis and chemoresistance of basal-like spheroid breast cancer model, that could be potentially used as therapeutic tools to improve the diagnosis, prognosis and treatment of triple negative breast cancers (Abboodi, 2014, Calcagno, Salcido et al., 2010). The enriched pathways inferred from the bioinformatics analysis which may be crucial in basal-like tumorigenesis are 'regulation of actin cytoskeleton' and 'focal adhesion'. These cytoskeleton remodelling signalling pathways targeted by miR-190b and miR-139 represent the important pathways at the early induction of EMT. The reorganisation of the cytoskeletal changes the cells morphologically and such changes permit the transformation to a motile mesenchymal state from the epithelial state. As a result, this transformation

that occurs during cancer progression enables the cancer cells to invade and migrate to other parts of the body (Sun, Sang et al., 2015).

Overall, both up- and down-regulated miRNAs demonstrated in spheroid-enriched MDA-MB-231 cells indicate higher cell proliferation, invasion and migration, metastasis, involvement of EMT, and maintenance self-renewal, and support the biological significance of the CSCs enrichment in the spheroid model of the present study. A deeper investigation for those miRNAs which have not been previously linked to breast cancer and had a fold change >10 in the present study is needed as these miRNAs could be potential candidates, both as marker or therapeutic target for CSCs-based research as the exact roles of these miRNAs have not been fully elucidated to date.

Table 6.12. The expression profiles of all the top ten miRNAs found in this study and matched to the existing literatures. Uniquely and commonly up- and down-regulated differentially expressed miRNAs in MDA-MB-231 spheroids of each cell type when compared to their parental counterparts

	miRNA	Deregulation	Tissue type	Validated target (s)	Mechanism	References
Uniquely expressed in MDA-MB-231 spheroids in relative to parental cells	**Breast CSCs-related miRNAs found in this study**					
	miR-24-1	Downregulated	Breast	FIH1	Drug resistance and sphere formation	(Roscigno, Puoti et al., 2017)
	miR-181a/b/c-1/2	Upregulated	Breast	ATM, YWHAG, PTEN	EMT and sphere formation	(Wang, Yu et al., 2011, Yoo, Kwak et al., 2016, Zhang and Zhang, 2015)
	miR-34c	Downregulated	Breast	Notch4	Self-renewal and EMT	(Yu, Jiao et al., 2012)
	miR-200a	Downregulated	Breast	ZEB1 and ZEB2	Self-renewal and EMT	(Feng, Wang et al., 2014, Sánchez-Cid, Pons et al., 2017)
	Breast cancer-related miRNAs found in this study					
	miR-133a-1	Upregulated	Breast	EGFR	Proliferation	(Cui, Zhang et al., 2013)
	miR-342	Upregulated	Breast	ID4	Tumorigenesis	(Crippa, Lusa et al., 2014)
	miR-205	Upregulated	Breast	ErbB3 and VEGF-A	Proliferation and sphere formation	(Wu, Zhu et al., 2009)
	miR-210	Upregulated	Breast	N/A	Proliferation and invasion	(Rothé , Ignatiadis et al., 2011)
	miR-129	Upregulated	Breast	Lamins	Tumorigenesis	(Setijono, Park et al., 2018)
	miR-139	Upregulated	Breast	Various	Cell motility and invasion	(Krishnan, Steptoe et

			genes in metastatic pathways		al., 2013)
	miR-769	Downregulated	Breast	NDRG1	Proliferation	(Luo, Chang et al., 2014)
	miR-204	Downregulated	Breast	FOXA1	Proliferation, migration and invasion	(Shen, Huang et al., 2017)
	miR-503	Downregulated	Breast	SMAD2	EMT	(Zhao, Fan et al., 2017)
	miR-33a	Downregulated	Breast	ADAM9 and ROS1	Stemness and metastasis	(Zhang, Zhang et al., 2015)
	miR-1271	Downregulated	Breast	SPIN1	Proliferation, migration and invasion	(Du and Liu, 2018)
	miR-802	Downregulated	Breast	FoxM1	Tumorigenesis	(Yuan and Wang, 2015)
	miR-190b	Downregulated	Breast	ZEB1	EMT	(Yu, Luo et al., 2018)

N/A; Not available

6.4.7 Common Cluster of Differentially Expressed miRNAs in Both MCF-7 and MDA-MB-231 Enriched CSCs Spheroid Models Possibly Associated with the Maintenance of CSCs Properties

To study the roles of miRNAs in the regulation of spheroid-enriched CSCs, the miRNA transcriptomic profiles in MCF-7 and MDA-MB-231 spheroid cells were identified and compared. Herein, we address a cluster of potential miRNAs associate with the CSCs-signalling pathways that could aid in the development of therapeutic strategies targeting miRNAs for future CSCs therapy.

Our study demonstrated that there were 20 significantly deregulated miRNAs that overlapped between these two spheroids models, suggesting that there were only a small number of miRNAs are being shared by the spheroid models. Of these 20 miRNAs, 10 were up-regulated and 10 were down-regulated.

Previous researches on miRNAs have demonstrated that some of the miRNAs with increased expression in cancer are normally oncogenic, meanwhile, those miRNAs with reduced expression in cancer are often associated with tumour suppressor roles. Although many studies have shown the crucial roles of miRNAs in the regulation of breast CSCs, studies on expression of miRNAs on CSCs-enriched spheroid models have been limited. Most of the targets were found to be involved in the regulation of various gene ontology related to biological processes such as signal transduction, cell proliferation and various transcriptional regulations. The enrichment analysis

of the commonly expressed miRNAs suggests that those miRNAs were involved in some crucial cancer-associated and also regulation of stem cells pathways.

We only discuss the top ten deregulated miRNAs which were commonly expressed in MCF-7 and MDA-MB-231 spheroid-enriched CSCs models. In the concept of CSCs, some pathways are responsible for the induction of EMT, maintaining stemness of CSCs, increase in proliferation, invasion and migration that distinguish stem cells from non-stem cells (Takebe, Miele et al., 2015). A few potential targeting pathways have been previously identified to eliminate CSCs in cancers, mainly focusing on altering the self-renewal mechanisms (Wnt, Notch, Hedgehog) or inhibiting the tumour progression and metastasis (JAK-STAT, TGF-β, PDFGR) (Pattabiraman and Weinberg, 2014). Apart from these common pathways, new therapies targeting the complex biology of CSCs were also being reported (Shibuya, Okada et al., 2015).

Here, we address a few potential miRNAs found in this study associated with the CSCs-signalling pathways that could aid the development of therapeutic strategies targeting miRNAs for CSC therapy.

6.4.8 Commonly Deregulated miRNAs Involved in BrCa CSCs

In the present study, deregulated miRNAs that have been linked to regulate CSCs in breast cancer only comprises miR-15b, miR-34a, miR-628, and miR-7 cluster, whereas the majority of the miRNAs were involved in the regulation of CSCs in other cancers which include those miRNAs listed in Table 6.13.

Mir-34a is one of the key players in tumour suppression and its occurrence is widely recognised in several types of cancers, suggesting its potential use for targeted therapies (Misso, Di Martino et al., 2014). MiR-34a was recently reported to regulate plasticity of epithelial cells in a model of human breast cancer. Their reduced expression provided evidence of promoting basal-like commitment and enhancing the CSCs pool by regulating the expression of genes associated with the Wnt signalling pathway (Bonetti, Climent et al., 2019). Reduced expression of miR-34a has been shown to enhance metastatic ability of lymphatic pancreatic cancer enriched with CSCs (Luo, Long et al., 2013). Similarly, the pronounced effects of miR-34a through its negative regulation on cell proliferation, invasion and migration have been observed in several types of cancer CSCs such as breast cancer (Kang, Sun et al., 2015), lung cancer (Shi, Au et al., 2014) and colon cancer (Tazawa, Tsuchiya et al., 2007). The experimentally validated target genes of miR-34a such as NOTCH1 and Bcl-2 reside mainly in the modulation of CSCs self-renewal and cell fate determination where it directly targets Notch receptors (Bu, Chen et al., 2013, Li, Ren et al., 2014). In fact, a wide range of regulation miR-34a on cancer progression and invasion have been documented such as in

the control of cell proliferation (Navarro and Lieberman, 2015), migration (Li, Yuan et al., 2013), immune evasion (Raisch J, Darfeuille-Michaud A et al., 2013), or apoptosis (Yamaguchi, Hashiguchi et al., 2008). Besides being involved in self-renewal pathway, miR-34a was also associated with MAPK signalling pathway by deregulating MAPK–related genes such as MAP2K1 (Ikeda, Tanji et al., 2012). Therefore, miR-34a which is important for the self-renewal properties of the spheroid-enriched CSCs via the Notch signalling pathway, may act as a prognosticator for breast CSCs.

On the other hand, another prominent down-regulated miRNAs that play important roles in regulating CSCs in breast and prostate cancers is miR-7-1//-2//-3. Restoration of miR-7 in breast CSCs was significantly found to reduce the ability for sphere formation, invasion and self-renewability via modulating KLF4 expression (Okuda, Xing et al., 2013). In a similar study, overexpression of miR-7 negatively regulates KLF4 expression, which promotes tumorigenesis in epithelial-derived cancer types (Meza-Sosa, EI. et al., 2014). Likewise, in a breast CSCs model, reversal of EMT and inhibition on invasion and metastasis was associated with increased miR-7 expression by targeting SETDB1 oncogene in the STAT3 pathway.

Another inducer of EMT is miR-15 cluster, which promote EMT in breast cancer by targeting self-renewal factors such as Bmi-1. These factors are important proteins that regulate self-renewal via repression of cluster of genes associated with senescence and apoptosis. It was reported that miR-15b

targeted this factor that enhanced the capacity for stem cell renewability not just in breast CSCs but also in normal progenitor cells (Shimono, Mukohyama et al., 2016). Our prediction indicates that VEGFA could be the potential target of miR-15b, one of the genes involved in VEGF signalling pathway. Previous studies indicated that miR-15b was involved in tumour angiogenesis, via regulating VEGF expression under hypoxic condition (Hua, Lv et al., 2006, Madanecki, Kapoor et al., 2013). Therefore, hypoxic condition in the core of spheroids (Riffle, Pandey et al., 2017) could possibly lead to the deregulation expression of miR-15b in spheroids relative to the monolayer cells. It was also reported that the anti-proliferative effect due to the elevation of miR-15a reduces stemness in CD44+ subpopulation. The overexpression was also reported in other study that sensitised the cells to chemotherapeutic drugs by activating ABCG2 (An, Sarmiento et al., 2016). Therefore, this suggests the potential use of miR-15b as a new therapeutic approach.

Taken together, both miR-34a and miR-7 found to be commonly expressed in the spheroid-enriched CSCs models were tumour suppressors and the overexpression of these miRNAs may provide new directions for treating higher invasive stage of breast cancers and also CSCs. In fact, advanced therapy utilising novel delivery system such as nanoparticles have been developed. MiR-34a encapsulated in a liposomal plasmid nanoparticle delivery system injected into established BrCa CSCs reduced cell proliferation *in vivo* and eventually eradicate the tumours, proving it to be the next potential targeted therapy (Lin, Chen et al., 2017).

6.4.9 Commonly Deregulated miRNAs Involved in Cancers and CSCs of Other Cancers

Notably, our findings also revealed a handful of deregulated miRNAs associated with CSCs and cancers of different tissues that have not previously been reported in breast CSCs (Table 6.13). Specifically, miR-148a, which has dual roles, either as tumour suppressor or oncogenic miRNAs, in various cancers including colorectal, pancreatic, breast and lung (Li, Deng et al., 2016). MiR-148a was apparently found to be down-regulated in most of the breast cancer studies (Xu, Jiang et al., 2013, Yu, Li et al., 2011), but remains largely unclear in the context of breast CSCs. Nevertheless, recent studies have demonstrated that miR-148a, a novel miRNAs, was found to suppress CSCs properties in hepatocellular carcinoma by targeting the Wnt signalling pathway (Mu, Zhu et al., 2017). In addition, overexpression of miR-148a was also found to regulate Hedgehog signalling pathway by directly targeting the GAS1 and LRP2 transcripts (Liu, He et al., 2015).

The potential targets of miR-628 were predicted to be involved in mTOR signalling pathway (Hua, Lv et al., 2006, Matter, Decaens et al., 2014, Wu, Yang et al., 2009). Studies have shown that inhibiting mTOR pathway reduces cell growth and tumour vascularity in hepatocellular carcinoma cell models (Matter, Decaens et al., 2014). The function of miR-196b in breast cancer is still uncertain.

Our study showed that miR-196b was down-regulated in both spheroid subtypes meanwhile other study showed that miR-196b was specifically up-regulated leading to reduction of apoptosis in colorectal cancer cells by mediating FAS gene expression (Mo, Alam et al., 2015). Similarly, overexpression of miR-196b was significantly contributed to maintenance of stemness and increase chemoresistance via regulation of STAT3 signalling pathway (Ren, Lin et al., 2017). The conflicting results might be attributed by the genetics of the different cell lines used and also the targeted genes which are yet to be validated.

In addition, miR-381 has been shown in regulating multidrug resistance in leukaemia cells by reducing expression of the drug efflux pump, and subsequently increasing the drug uptake by the cells (van Schooneveld, Wouters et al., 2012). MiR-411 was also recently being reported to target ITCH gene, which promotes cell proliferation and anchorage-independent growth by negative regulation in hepatocellular carcinoma (Xia, Zhang et al., 2015). This is also in agreement with our study where the spheroids had enhanced cell proliferation in the serum-free condition compared to the dissociated parental cells. Elevated expression of miR-127, a miRNA observed in embryo lung development, was found to promote EMT enabling the lung cancer cells to acquire the stem cell-like properties (Shi, Wang et al., 2017). These three miRNAs (miR-381, miR-411 and miR-127), found to be up-regulated in spheroids were down-regulated in other studies (Ma, 2010, Tan, Zhang et al., 2009, van Schooneveld, Wouters et al., 2012). This discrepancy is clear as miRNA profiles are attributed to the different experimental setting

such as source of the cell type used in the studies. Our study compared the miRNAs expression of the spheroid-enriched CSCs cells to that of the parental cells while other studies measured the miRNA expression in breast cancer tissues. Therefore, the profiles of spheroid-enriched CSCs cells and non-CSC are not comparable with each other.

Collectively, the miRNA-NGS data have been validated with the real-time PCR assay thus supporting the use and interpretation of miRNA expression profiles, inferred target genes and enriched pathways based on the bioinformatics data analysis. It is worth exploring the other afore-mentioned miRNAs that are involved in the other stem cell-related signalling pathways. Therefore, we hypothesised that this cluster of deregulated miRNAs plays important roles in the regulation of CSCs in the two subtypes of breast cancer and perhaps in other cancers. Further studies involving knock-out or over expression can be performed to further verify the authenticity interactions of miRNA-target genes. Modulation of these miRNAs could be a promising approach to suppress those pathways, thus improving the effectiveness of breast cancer treatments.

Table 6.13. The expression profiles of all the top ten miRNAs found in this study and matched to the existing literatures. Common up- and down-regulated differentially expressed miRNAs found in both MDA-MB-231 and MCF-7 spheroids

	miRNA	Deregulation	Tissue type	Validated target (s)	Mechanism	References
Commonly expressed in both spheroid models	**Breast CSCs-related miRNAs found in this study**					
	miR-15b	Downregulated	Breast	BMI1	Maintain self-renewal capacity Promote MET to EMT	(Patel, Garikapati et al., 2018, Shimono, Mukohyama et al., 2016)
	miR-34a	Downregulated	Breast	NOTCH1	Regulate CSCs-associated pathways	(Kang, Mao et al., 2015)
	miR-628	Downregulated	Breast	SOS1	Promote migration and invasion	(Lin, Gao et al., 2018)
	miR-7-1//-2//-3	Downregulated	Breast	SETDB1	Promote invasion and EMT	(Zhang, Cai et al., 2014)
	Other CSCs-related miRNAs found in this study					
	miR-1246	Upregulated	Pancreas	N/A	Enhance chemoresistance and stemness	(Xu, Hannafon et al., 2017)
	miR-1246 miR-127 miR-410	Upregulated	Lung	N/A N/A Gsk3β	Enhance CSCs properties Promote EMT	(Chen, Wu et al., 2018, Ke, Yuan et al., 2017, Kim, An et al., 2016, Shi, Wang et al., 2017, Xu, Luo et al., 2016)
	miR-18a, miR-191	Downregulated		ATM and HIF-1α HIF-2α		

miR-126	Upregulated	Leukemia	N/A	Promote self-renewal	(Lechman, Gentner et al., 2016)
miR-143 miR-145 miiR-34a miR-7-1//-2//-3	Upregulated Downregulated	Prostate	FNDC3B N/A CD44 KLF4/PI3K/Akt/p21	Promote tumorigenesis, regulate CSCs characteristics Promote tumour regeneration and metastasis	(Chang, Zhou et al., 2015, Fan, Chen et al., 2013, Huang, Guo et al., 2012, Liu, Kelnar et al., 2011)
miR-145	Upregulated	Cervical	N/A	Regulate core CSCs transcription factors	(Zhou, Yue et al., 2017)
miR-148a	Downregulated	Liver	TGF-ß	Promote sphere formation and stemness	(Jiang, Mu et al., 2014)
miR-148a	Downregulated	Thyroid	INO80	Promote sphere formation and stemness	(Sheng, Chen et al., 2016)
miR-15b miR-191 miR-3613	Downregulated	Oral	TRIM14 N/A N/A	Promote stemness and chemoresistance Biomarker	(Liu, Ma et al., 2014, Wang, Guo et al., 2017)
miR-196b	Downregulated	Colorectal	STAT3	Promote stemness and chemoresistance	(Ren, Lin et al., 2017)
miR-3613 miR-628	Downregulated	Ovarian	N/A FGFR2	Biomarker Promote stemness	(Chong, Jeon et al., 2015, Li, Qian et al., 2018)
Other cancers-related miRNAs found in this study					
miR-1246	Upregulated	Breast	CCNG2	Promote tumorigenesis and	(Guo, Yuan

miR-211 miR-411			N/A N/A	chemoresistance	et al., 2016, Lee, Lee et al., 2016, Li, Ren et al., 2017)
miR-126 miR-127 miR-145	Upregulated	Breast	N/A BCL6 N/A	Biomarker Exosomal biomarker Prognostic biomarker	(Bhat, Majid et al., 2019, Chen, Wang et al., 2013, Quan, Huang et al., 2018)
miR-1291	Upregulated	Pancreatic	N/A	Increase multi-drug resistance	(To, 2013)
miR-211	Upregulated	Melanoma	N/A	Inhibit invasion	(Mazar, Qi et al., 2016)
miR-381	Upregulated	Breast	Twist1	Inhibit EMT	(He, Wei et al., 2016)
miR-411 miR-3074	Upregulated Downregulated	Prostate	N/A N/A	Biomarker Inhibit cell proliferation	(Gu, Shi et al., 2018, Nam, Amemiya et al., 2015)
miR-15b	Downregulated	Colorectal	NF-κB	Promote chemoresistance	(Zhao, Zhao et al., 2017)
miR-628	Downregulated	Various	FGFR2	Biomarkers	(Li, Sun et al., 2018)

N/A: Not available

6.5 Current Limitations

A possible limitation of this study is that the spheroid enriched CSCs study models were derived from only two breast cancer cell lines for the analysis of the miRNA expression profiles. Therefore, it remains unknown whether the similar observation would be seen if spheroid enriched CSCs were derived from the breast cancer tissues, specifically on the miRNAs differential expression patterns. It has been reported that the different culture conditions that exist between cells in *in vitro* and *in vivo* environments pose technical challenges in establishing the relevant models for studying breast cancer (Cope, Fackler et al., 2014). On top of that, it is still questionable as on how representative cell lines are as they could not fully capture the tumour heterogeneities found in patient-derived tissues. A number of breast cancer cell lines have demonstrated accumulation of genetic aberrations especially those derived from highly invasive tumours especially when they were subjected to long term *in vitro* cultures (Dai, Cheng et al., 2017). One of the highly controversial cell lines in breast cancer research is the authentication of MDA-MB-435 which was initially thought to be originated from a breast carcinoma, but then turned out to be melanoma origin (Korch, Hall et al., 2018). Despite controversies that surrounding the use of breast cancer cell lines in research, cell lines are still considered to be feasible models for tumours. Systematic analyses via modern molecular profiling techniques have enabled the classification of breast cancer cell lines into their different subtypes and ranked them for their first choice of experimental usage (Kao, Keyan Salari et al., 2009). Thus, together with the past and improved knowledge on breast cancer cell lines gained through advances in molecular profiling, numerous studies

have supported the use of MCF-7 and MDA-MB-231 breast cancer cell lines as tumour models (Holliday and Speirs, 2011). Thereby, the use of these two cell lines as the *in vitro* model in the present study deemed to be useful for the understanding of CSCs mechanisms.

Our results indicate that dysregulation of miRNAs play an important role to differentiate the CSCs subpopulation in both spheroid subtypes. These panel of dysregulated miRNAs may be responsible for maintaining the stem cell-like properties of breast CSCs. It is worth noting that the possibility to further analyses the miRNAs datasets using The Cancer Genome Atlas Program (TCGS) and (Genomic Spatial Event) GSE databases to validate the potential miRNAs as prognostics or diagnostic markers in breast cancer carcinogenesis.

6.6 Summary

To summarize, this work is one of the first to provide a comprehensive comparison in terms of the miRNAs expressions in the spheroid-enriched CSCs model. The work described here demonstrated the biological features and the roles played by certain miRNAs of the spheroid-enriched CSCs model. Interestingly, the gene annotation and pathways inferred from the uniquely expressed cluster of miRNAs found in the two breast spheroids models derived from the two breast cancer subtypes were similar. This indicates the highly complex interactions of the different groups of miRNAs in maintaining the CSCs and also their biological properties in different subtypes of breast cancer. It is also interesting to note that MCF-7 and MDA-MB-231 spheroid cells may

use different miRNAs to regulate their biological properties, but those miRNAs converged to the similar pathways and gene products. The different miRNAs profiles are explained by the fact that MCF-7 and MDA-MB-231 were two different kinds of breast cancer cell lines, representing luminal and basal-like breast cancer subtypes, respectively. The differences may also be attributed by the status of their EMT, metastasis and their chemoresistance properties. Besides, miRNAs play important roles in the expression of CSCs characteristics. A set of new possibly new miRNA signatures found commonly expressed in both spheroid-enriched CSCs model demonstrated in this study may contribute to new biomarker discovery, and new diagnosis, prognosis and treatment options.

CHAPTER 7

CONCLUSIONS AND FUTURE STUDIES

7.1 Conclusion

In conclusion, we have demonstrated that the spheroid culturing method can be used to enrich for CSCs-like subpopulations in both breast cancer cell lines as shown in the present study. Overall, these data suggested that the spheroid cells generated from the two most common breast cancer cell lines were phenotypically different particularly in their CSC-like characteristics. Both spheroid cells demonstrated increased in self-renewability, expression of stem cell- and CSCs-related markers, possessed higher cell proliferation, exhibited proliferation, migration and wound healing abilities, and demonstrated a higher resistant against tested chemotherapy drugs relative to their parental cells. However, higher proportions of CSCs subpopulations were enriched in MDA-MB-231 spheroid compared to the MCF-7 spheroids, indicated by their increased ALDH activity.

Our results also showed that both spheroids expressed certain miRNAs unique to their cancer subtypes, and also a cluster of miRNAs that were commonly expressed in both spheroids. This study not only provided further understanding on the roles of miRNAs in spheroid-enriched CSCs models, but more importantly highlighting certain miRNAs that could possibly to be used to differentiate the spheroid subtypes. The subsets of exclusively expressed miRNAs generally enriched unique target genes but modulate a few similar pathways, suggesting the highly complex miRNAs-target genes interactions

that focus on maintaining breast cancer carcinogenesis. These exclusive miRNAs may be potential diagnostic or prognostics markers for the respective disease subtype. Based on our findings, the 20 commonly deregulated miRNAs found in both spheroids cells might be potential targets in breast CSCs-related studies**.** In addition, the research also points towards some of the cluster of miRNAs and their predicted target genes and pathways that could be utilised as potential interventional approaches for CSCs control. Nevertheless, much remains to be explored about the involvement of these miRNAs in the events of breast CSCs, for instance, either as onco-miRNAs or tumour suppressor miRNAs.

7.2 Future Studies

In this study, the differentially expressed miRNA profiles, functional annotations, target genes network interactions and pathways uniquely expressed in each of the spheroid cells and those shared by both spheroid cells were highlighted. Our miRNA transcriptomic profiling suggested both of the spheroid-enriched CSCs derived from both cell lines may regulate different group of miRNAs, which may represent the distinct subtypes of breast cancer. Therefore, it would be interesting to explore the contribution of the different group of miRNAs in the breast cancer subtype classification that may help to develop potential biomarkers for breast cancer prognostication. Interestingly, some pathways associated with the regulation of CSCs were identified from the predicted targets of the overlapped miRNAs, and other pathways involved in the regulation of hormonal level and metastasis mechanism were also found in the other group of miRNAs. Notably, the predicted pathways by the 20

commonly deregulated miRNAs found in both spheroid cells through bioinformatics analysis revealed two important new pathways, i. mTOR signalling pathway and ii. Focal adhesion pathway, which has recently generated great interest. mTOR signalling pathway was reported as the positive master regulator of cell growth and proliferation in many cancers (Laplante and Sabatini, 2009, Paquette, El-Houjeiri et al., 2018). More specifically, closer examination of mTOR pathway showed the extensive cell proliferation was due to the activation of a group of antiapoptotic proteins that ultimately lead to tumour progression (Pópulo, Lopes et al., 2012). Therefore, direct suppression of mTOR pathway via the targeted genes through regulation of the specific miRNAs could be a next promising strategy to target cancer (Advani, 2010). On the other hand, cell adhesion molecules such as cadherins and integrins involved in focal adhesion pathway are important to regulate interaction between cell to cell and also cell to the extracellular matrices (ECM) (Farahani, Patra et al., 2014). CSCs have been known to maintain its stemness and ability to undergo EMT through the changes in this group of molecules and focal adhesion kinases signalling (Begum, Ewachiw et al., 2017, Schober and Fuchs, 2011, Zhou, Yi et al., 2019). Loss of focal adhesion mediated signalling pathway lead to disruption of cell-ECM interactions leading to cell anoiksis and resulted in cell death (Paoli, Giannoni et al., 2013). As a result, lack of such signals could contribute to detachment-induced cell death, one of the crucial steps in inhibiting tumour progression and metastasis. In summary, future studies could be directed towards these two important pathways to facilitate breast cancer management.

The present study only assessed miRNA transcriptomic of the spheroid cells derived from the breast cancer cell lines, which are most likely consisted of a heterogeneous cell population. Two well-accepted breast CSCs surface markers, namely CD44 and CD24 have been used extensively to isolate CSCs from solid breast tumours (Ricardo, Vieira et al., 2011). As few as 100 cells, CD44+ cells promoted tumorigenesis in breast cancer displaying stem cell properties such as self-renewal and regulation of EMT (Fiori, Di Franco et al., 2019, Sheridan, Kishimoto et al., 2006, Yan, Chen et al., 2013). Recent evidences indicated that breast CSCs are phenotypically diverse and there exist many subsets of breast CSCs within the tumour (Jaggupilli and Elkord, 2012). As a consequence of the proposed difference in the subset of the breast CSCs, we sought to investigate this by miRNA profiling of the different sorted cell populations within our spheroid cell models. Through our initial investigations, our preliminary findings demonstrated that the double negative cell population (CD44-/CD24-) could also form spheroids and resulted in a higher sphere-forming efficiency percentage. This next phase of study is hoped to increase our understanding of CSCs function in tumour heterogeneity and the importance of miRNA regulation in the CSCs subpopulations. Moreover, the establishment of the miRNA expression profiles could help to clarify the significance of the surface markers for breast CSCs. Establishment of the expression profiles could help to reveal several miRNAs that are underexpressed or overexpressed among breast cancers as well as among the different subpopulations, thereby identifying miRNAs that could improve our understanding of breast cancer carcinogenesis.

References

Abboodi, M., 2014. Isolation, identification, and spheroids formation of breast cancer stem cells, therapeutics implications. *Clinical Cancer Investigation Journal,* 3(4) pp. 322 - 325.

Abdullah NA et al., 2013. Survival rate of breast cancer patients in Malaysia: a population-based study. *Asian Pacific Journal of Cancer Prevention (APJCP),* 14(8) pp. 4591 - 4594.

Abel, E.V. et al., 2014. The Notch pathway is important in maintaining the cancer stem cell population in pancreatic cancer. *PLoS ONE,* 9(3) pp. e91983.

Acunzo, M., Romano, G., Wernicke, D. and Croce, C.M., 2015. MicroRNA and cancer - A brief overview. *Advances in Biological Regulation,* 57 pp. 1 - 9.

Advani, S., 2010. Targeting mTOR pathway: A new concept in cancer therapy. *Indian Journal of Medical and Paediatric Oncology,* 31(4) pp. 132 - 136.

Agarwal, A., Mulgund, A., Hamada, A. and Chyatte, M.R., 2015. A unique view on male infertility around the globe. *Reproductive Biology and Endocrinology,* 13(1) pp. e37.

Ahmed, K. et al., 2019. Towards the first targeted therapy for triple-negative breast cancer: Repositioning of clofazimine as a chemotherapy-compatible selective Wnt pathway inhibitor. *Cancer Letters,* 449 pp. 45 - 55.

Ahmed, N. et al., 2018. Tumour microenvironment and metabolic plasticity in cancer and cancer stem cells: Perspectives on metabolic and immune regulatory signatures in chemoresistant ovarian cancer stem cells. *Seminars in Cancer Biology,* 53 pp. 265 - 281.

Ahn, S.-H. et al., 2018. Effect of necrosis on the miRNA-mRNA regulatory network in CRT-MG human astroglioma cells. *Cancer research and treatment : official journal of Korean Cancer Association,* 50(2) pp. 382 - 397.

Aisner, D.L. et al., 2018. The impact of smoking and TP53 mutations in lung adenocarcinoma patients with targetable mutations-the lung cancer mutation consortium (LCMC2). *Clinical Cancer Research,* 24(5) pp. 1038 - 1047.

Akhavantabasi, S., Sapmaz, A., Tuna, S. and Erson-Bensan, A.E., 2012. miR-125b targets ARID3B in breast cancer cells. *Cell Structure and Function,* 37(1) pp. 27 - 38.

Akrap, N. et al., 2016. Identification of distinct breast cancer stem cell populations based on single-cell analyses of functionally enriched stem and progenitor pools. *Stem Cell Reports,* 6(1) pp. 121 - 136.

Al-Hajj, M. et al., 2003. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the USA,* 100(7) pp. 3983 - 3988.

Al-Mahmood, S., Sapiezynski, J., Garbuzenko, O.B. and Minko, T., 2018. Metastatic and triple-negative breast cancer: challenges and treatment options. *Drug Delivery and Translational Research,* 8(5) pp. 1483 - 1507.

Alvarez M.L., 2014. Isolation of urinary exosomes for RNA biomarker discovery using a simple, fast, and highly scalable method. In: Alvarez M., and Nourbakhsh M. (eds). RNA Mapping. Methods in Molecular Biology (Methods and Protocols), vol 1182. Humana Press, New York, NY, pp. 145 – 170.

An, X., Sarmiento, C., Tan, T. and Zhu, H., 2016. Regulation of multidrug resistance by microRNAs in anti-cancer therapy. *Acta pharmaceutica Sinica. B,* 7(1) pp. 38 - 51.

Anampa, J., Makower, D. and Sparano, J.A., 2015. Progress in adjuvant chemotherapy for breast cancer: an overview. *BioMed Central Medicine,* 13(1) pp. 195.

Anand, S. et al., 2010. MicroRNA-132-mediated loss of p120RasGAP activates the endothelium to facilitate pathological angiogenesis. *Nature medicine,* 16(8) pp. 909 - 914.

Andorfer, C.A., Necela, B.M., Thompson, E.A. and Perez, E.A., 2011. MicroRNA signatures: clinical biomarkers for the diagnosis and treatment of breast cancer. *Trends in Molecular Medicine,* 17(6) pp. 313 - 319.

Antoni, D., Burckel, H., Josset, E. and Noel, G., 2015. Three-dimensional cell culture: a breakthrough in vivo. *International Journal of Molecular Sciences,* 16(3) pp. 5517 - 5527.

Apostolou, P. and Fostira, F., 2013. Hereditary breast cancer: the era of new susceptibility genes. *BioMed Research International,* 2013 pp. 747318

Arab-Bafrani, Z. et al., 2016. Culturing in serum-free culture medium on collagen type-I-coated plate increases expression of CD133 and retains original phenotype of HT-29 cancer stem cell. *Advanced Biomedical Research,* 5(59).

Asnarulkhadi, A.S., Maryam, A. and Latiffah A., L., 2016. Insufficient knowledge of breast cancer risk factors among Malaysian female university students. *Global Journal of Health Science,* 8(1) pp. 277 - 285.

Avgeris, M., Stravodimos, K. and Scorilas, A., 2014. Loss of miR-378 in prostate cancer, a common regulator of KLK2 and KLK4, correlates with aggressive disease phenotype and predicts the short-term relapse of the patients. *Biological Chemistry*, 395(9) pp. 1095 - 1104.

Ayob, A.Z. and Ramasamy, T.S., 2018. Cancer stem cells as key drivers of tumour progression. *Journal of Biomedical Science,* 25(1) pp. 20.

Azizah Ab M et al. 2016. *National Cancer Registry Report 2007-2011 (Malaysian Cancer Statistics - Data and Figure)* National Cancer Institute.

Bagheri, V. et al., 2018. Isolation and identification of chemotherapy-enriched sphere-forming cells from a patient with gastric cancer. *Journal of Cellular Physiology,* 233(10) pp. 7036 - 7046.

Bahmad, H.F. et al., 2018. Sphere-formation assay: three-dimensional *in vitro* culturing of prostate cancer stem/progenitor sphere-forming cells. *Frontiers in Oncology*, 8 pp. 347.

Bailey, M.H. et al., 2018. Comprehensive characterization of cancer driver genes and mutations. *Cell,* 173(2) pp. 371 - 385.

Bailey, S.T., Westerling, T. and Brown, M., 2015. Loss of estrogen-regulated microRNA expression increases HER2 signaling and is prognostic of poor outcome in luminal breast cancer. *Cancer Research,* 75(2) pp. 436 - 445.

Balacescu, O. et al., 2018. The impact of miRNA in colorectal cancer progression and its liver metastases. *International Journal of Molecular Sciences*, 19(12) pp. 3711.

Balatti, V. et al., 2018. miR-125a and miR-34a expression predicts Richter syndrome in chronic lymphocytic leukemia patients. *Blood,* 132(20) pp. 2179.

Bao, B. et al., 2013. Overview of cancer stem cells (CSCs) and mechanisms of their regulation: implications for cancer therapy. *Current protocols in pharmacology*, Chapter 14 pp. Unit-14.25.

Bar, I. et al., 2017. The microRNA miR-210 is expressed by cancer cells but also by the tumor microenvironment in triple-negative breast cancer. *Journal of Histochemistry and Cytochemistry,* 65(6) pp. 335 - 346.

Bartakova, A. et al., 2017. CD44 as a cancer stem cell marker and its prognostic value in patients with ovarian carcinoma. *Journal of Obstetrics and Gynaecology,* 38(1) pp. 110 - 114.

Basu, A.K., 2018. DNA damage, mutagenesis and cancer. *International Journal of Molecular Sciences,* 19(4) pp. 970.

Begley, C.G. and Ellis, L.M., 2012. Raise standards for preclinical cancer research. *Nature,* 483 pp. 531 - 533.

Begum, A. et al., 2017. The extracellular matrix and focal adhesion kinase signaling regulate cancer stem cell function in pancreatic ductal adenocarcinoma. *PLoS ONE,* 12(7) pp. e0180181.

Beumer, I. et al., 2016. Equivalence of MammaPrint array types in clinical trials and diagnostics. *Breast Cancer Research and Treatment,* 156(2) pp. 279 - 287.

Bhat, S.A., Majid, S. and Hassan, T., 2019. MicroRNAs and its emerging role as breast cancer diagnostic marker- A review. *Advances in Biomarker Sciences and Technology,* 1 pp. 1 - 8.

Bianchini, G. et al., 2016. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nature Reviews Clinical Oncology,* 13(11) pp. 674 - 690.

Biankin, A.V. et al., 2012. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature,* 491(7424) pp. 399 - 405.

Bisgin, H., Gong, B., Wang, Y. and Tong, W., 2018. Evaluation of bioinformatics approaches for Next-Generation Sequencing analysis of microRNAs with a toxicogenomics study design. *Frontiers in Genetics,* 9(22) pp. 29467792.

Blasco, M.A., 2005. Telomeres and human disease: ageing, cancer and beyond. *Nature Reviews Genetics,* 6(8) pp. 611 - 622.

Blenkiron, C. et al., 2007. MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biology,* 8(10) pp. R214.

Blume, R. et al., 2018. The molecular signature of AML with increased ALDH activity suggests a stem cell origin. *Leukemia and Lymphoma,* 59(9) pp. 2201 - 2210.

Bockhorn, J. et al., 2013. MicroRNA-30c inhibits human breast tumour chemotherapy resistance by regulating TWF1 and IL-11. *Nature Communications,* 4 pp. e1393.

Bomken, S., Fiser, K., Heidenreich, O. and Vormoor, J., 2010. Understanding the cancer stem cell. *British Journal of Cancer,* 103(4) pp. 439 - 445.

Bonetti, P. et al., 2019. Dual role for miR-34a in the control of early progenitor proliferation and commitment in the mammary gland and in breast cancer. *Oncogene,* 38(3) pp. 360 - 374.

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

Brabletz, T., Kalluri, R., Nieto, M.A. and Weinberg, R.A., 2018. EMT in cancer. *Nature Reviews Cancer,* 18 pp. 128.

Branch, P. et al., 2000. Spontaneous development of drug resistance: mismatch repair and p53 defects in resistance to cisplatin in human tumor cells. *Oncogene,* 19(28) pp. 3138 - 3145.

Bray, F. et al., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians,* 68(6) pp. 394 - 424.

Broadley, K.W.R. et al., 2011. Side population is not necessary or sufficient for a cancer stem cell phenotype in glioblastoma multiforme. *STEM CELLS,* 29(3) pp. 452 - 461.

Bu, P. et al., 2013. A microRNA miR-34a regulated bimodal switch targets Notch in colon cancer stem cells. *Cell stem cell,* 12(5) pp. 602 - 615.

Buermans, H.P.J. et al., 2010. New methods for next generation sequencing based microRNA expression profiling. *BioMed Central Genomics,* 11(1) pp. e716.

Burdall, S.E., Hanby, A.M., Lansdown, M.R. and Speirs, V., 2003. Breast cancer cell lines: friend or foe? *Breast Cancer Research,* 5(2) pp. 89 - 95.

Burkert, J., Otto, W.R. and Wright, N.A., 2008. Side populations of gastrointestinal cancers are not enriched in stem cells. *The Journal of Pathology,* 214(5) pp. 564 - 573.

Cai, J. et al., 2013. MicroRNA-374a activates Wnt/β-catenin signaling to promote breast cancer metastasis. *The Journal of Clinical Investigation,* 123(2) pp. 566 - 579.

Cai, Z. et al., 2018. Signalling mechanism(s) of epithelial-mesenchymal transition and cancer stem cells in tumour therapeutic resistance. *Clinica Chimica Acta,* 483 pp. 156 - 163.

Calcagno, A.M. et al., 2010. Prolonged drug selection of breast cancer cells and enrichment of cancer stem cell characteristics. *Journal of the National Cancer Institute,* 102(21) pp. 1637 - 1652.

Calin, G.A. et al., 2002. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences,* 99(24) pp. 15524 - 15529.

Calin, G.A. et al., 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America,* 101(9) pp. 2999 - 3004.

Cao, L. et al., 2011. Sphere-forming cell subpopulations with cancer stem cell properties in human hepatoma cell lines. *BMC Gastroenterology,* 11(1) pp. 71 - 82.

Capp, J.-P., 2019. Cancer stem cells: from historical roots to a new perspective. *Journal of Oncology,* 2019 pp. e10.

Carver, K., Ming, X. and Juliano, R.L., 2014. Multicellular tumor spheroids as a model for assessing delivery of oligonucleotides in three dimensions. *Molecular Therapy Nucleic Acids,* 3 pp. e153.

Cascione, L. et al., 2013. Integrated microRNA and mRNA signatures associated with survival in triple negative breast cancer. *PLoS ONE,* 8(2) pp. e55910.

Castro, M.A.A. et al., 2015. Regulators of genetic risk of breast cancer identified by integrative network analysis. *Nature Genetics,* 48(1) pp. 12 - 21.

Catalanotto, C., Cogoni, C. and Zardo, G., 2016. MicroRNA in control of gene expression: An overview of nuclea functions. *International Journal of Molecular Sciences,* 17(10) pp. e1712.

Catuogno, S. et al., 2011. Recent advance in biosensors for microRNAs detection in cancer. *Cancers,* 3(2) pp. 1877 - 1898.

Cebola, I.s. and Peinado, M.A., 2012. Epigenetic deregulation of the COX pathway in cancer. *Progress in Lipid Research,* 51(4) pp. 301 - 313.

Cha, S.Y. et al., 2017. Clinical Impact of microRNAs associated with cancer stem cells as a prognostic factor in ovarian carcinoma. *Journal of Cancer,* 8(17) pp. 3538 - 3547.

Chan, C.H.F., Coopey, S.B., Freer, P.E. and Hughes, K.S., 2015. False-negative rate of combined mammography and ultrasound for women with palpable breast masses. *Breast Cancer Research and Treatment,* 153(3) pp. 699 - 702.

Chan, Y.C., Banerjee, J., Choi, S.Y. and Sen, C.K., 2012. miR-210: the master hypoxamir. *Microcirculation (New York, N.Y. : 1994),* 19(3) pp. 215 - 223.

Chandler, J.M. and Lagasse, E., 2010. Cancerous stem cells: deviant stem cells with cancer-causing misbehavior. *Stem Cell Research and Therapy,* 1(2) pp. 13 - 22.

Chang, C.-W., Lo, J.-F. and Wang, X.W., 2019. Roles of mitochondria in liver cancer stem cells. *Differentiation,* 107 pp. 35 - 41.

Chang, J.C., 2016. Cancer stem cells: Role in tumor growth, recurrence, metastasis, and treatment resistance. *Medicine,* 95(Suppl 1) pp. S20 - S25.

Chang, Y.-L. et al., 2015. MicroRNA-7 inhibits the stemness of prostate cancer stem-like cells and tumorigenesis by repressing KLF4/PI3K/Akt/p21 pathway. *Oncotarget,* 6(27) pp. 24017 - 24031.

Chen, C., Zhao, S., Karnad, A. and Freeman, J.W., 2018. The biology and role of CD44 in cancer progression: therapeutic implications. *Journal of Hematology and Oncology,* 11(1) pp. 64.

Chen, D. et al., 2018. miR-27b-3p inhibits proliferation and potentially reverses multi-chemoresistance by targeting CBLB/GRB2 in breast cancer cells. *Cell Death and Disease,* 9(2) pp. 188.

Chen, J. et al., 2013. miR-127 regulates cell proliferation and senescence by targeting BCL6. *PLoS ONE,* 8(11) pp. e80266.

Chen, K., Huang, Y.-h. and Chen, J.-l., 2013. Understanding and targeting cancer stem cells: therapeutic implications and challenges. *Acta Pharmacologica Sinica,* 34(6) pp. 732 - 740.

Chen, P. et al., 2013. miR-9 is an essential oncogenic microRNA specifically overexpressed in mixed lineage leukemia-rearranged leukemia. *Proceedings of the National Academy of Sciences of the United States of America,* 110(28) pp. 11511 - 11516.

Chen, S.-F. et al., 2012. Nonadhesive culture system as a model of rapid sphere formation with cancer stem cell properties. *PLoS ONE,* 7(2) pp. e31864.

Chen, X. et al., 2018. Radiosensitizing effects of miR-18a-5p on lung cancer stem-like cells via downregulating both ATM and HIF-1α. *Cancer medicine,* 7(8) pp. 3834 - 3847.

Chien, W.W. et al., 2011. Cyclin-dependent kinase 1 expression is inhibited by p16INK4a at the post-transcriptional level through the microRNA pathway. *Oncogene,* 30(16) pp. 1880 - 1891.

Chiesa, F. and Sacchini, V.S., 2016. Risk-reducing mastectomy. *Minerva Ginecologica,* 68(5) pp. 544 - 547.

Chong, G.O. et al., 2015. Differential microRNA expression profiles in primary and recurrent epithelial ovarian cancer. *Anticancer Research,* 35(5) pp. 2611 - 2617.

Christopoulos, P.F., Msaouel, P. and Koutsilieris, M., 2015. The role of the insulin-like growth factor-1 system in breast cancer. *Molecular Cancer,* 14 pp. e43.

Cicatiello, L. et al., 2009. Estrogen receptor α controls a gene network in luminal-like breast cancer cells comprising multiple transcription factors and microRNAs. *The American Journal of Pathology,* 176(5) pp. 2113 - 2130.

Cochrane, C.R., Szczepny, A., Watkins, D.N. and Cain, J.E., 2015. Hedgehog signaling in the maintenance of cancer stem cells. *Cancers,* 7(3) pp. 1554 - 1585.

Coenen-Stass, A.M.L. et al., 2018. Evaluation of methodologies for microRNA biomarker detection by next generation sequencing. *RNA biology,* 15(8) pp. 1133 - 1145.

Collins, A.T. et al., 2005. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Research,* 65 pp. 10946 - 10951.

Cope, L.M. et al., 2014. Do breast cancer cell lines provide a relevant model of the patient tumor methylome? *PLoS ONE,* 9(8) pp. e105545.

Crippa, E. et al., 2014. miR-342 regulates BRCA1 expression through modulation of ID4 in breast cancer. *PLoS ONE,* 9(1) pp. e87039.

Cui, W. et al., 2013. microRNA-133a regulates the cell cycle and proliferation of breast cancer cells by targeting epidermal growth factor receptor through the EGFR/Akt signaling pathway. *Federation of European Biochemical Societies Letters,* 280(16) pp. 3962 - 3974.

Cui, X., Hartanto, Y. and Zhang, H., 2016. Advances in multicellular spheroids formation. *Journal of The Royal Society Interface,* 14(127) pp. 20160877.

Czerwinska, P. and Kaminska, B., 2015. Review: Regulation of breast cancer stem cell features. *Contemporary Oncology,* 19(1A) pp. 7 - 15.

Dahal Lamichane, B. et al., 2019. AGR2 is a target of canonical Wnt/β-catenin signaling and is important for stemness maintenance in colorectal cancer stem cells. *Biochemical and Biophysical Research Communications,* 515(4) pp. 600 - 606.

Dai, X., Cheng, H., Bai, Z. and Li, J., 2017. Breast cancer cell line classification and its relevance with breast tumor subtyping. *Journal of Cancer,* 8(16) pp. 3131 - 3141.

Dai, X. et al., 2015. Breast cancer intrinsic subtype classification, clinical use and future trends. *American Journal of Cancer Research,* 5(10) pp. 2929 - 2943.

Dar, A.A. et al., 2011. miRNA-205 suppresses melanoma cell proliferation and induces senescence via regulation of E2F1 protein. *The Journal of Biological Chemistry,* 286(19) pp. 16606 - 16614.

Dard-Dascot, C. et al., 2018. Systematic comparison of small RNA library preparation protocols for next-generation sequencing. *BioMed Central Genomics,* 19(1) pp. 118.

Das, P.M. and Singal, R., 2004. DNA methylation and cancer. *Journal of Clinical Oncology,* 22(22) pp. 4632 - 4642.

De Angelis ML et al., 2018. Colorectal cancer spheroid biobanks: multi-level approaches to drug sensitivity studies. *Cell Biology and Toxicology,* 34(6) pp. 459 - 469.

De Angelis, M.L., Francescangeli, F., La Torre, F. and Zeuner, A., 2019. Stem cell plasticity and dormancy in the development of cancer therapy resistance. *Frontiers in Oncology,* 9 pp. e626.

De Capoa, A. et al., 2003. DNA demethylation is directly related to tumour progression: Evidence in normal, pre-malignant and malignant cells from uterine cervix samples. *Oncology Reports,* 10(3) pp. 545 - 549.

De Francesco, E.M., Sotgia, F. and Lisanti, M.P., 2018. Cancer stem cells (CSCs): metabolic strategies for their identification and eradication. *The Biochemical journal,* 475(9) pp. 1611 - 1634.

Deb, B., Uddin, A. and Chakraborty, S., 2017. miRNAs and ovarian cancer: An overview. *Journal of Cellular Physiology,* 233(5) pp. 3846 - 3854.

Dehghan Harati, M., Rodemann, H.P. and Toulany, M., 2019. Nanog signaling mediates radioresistance in ALDH-positive breast cancer cells. *International Journal of Molecular Sciences,* 20(5) pp. e1151.

Dembinski, J.L. and Krauss, S., 2009. Characterization and functional analysis of a slow cycling stem cell-like subpopulation in pancreas adenocarcinoma. *Clinical and Experimental Metastasis,* 26(7) pp. 611 - 623.

Denkert, C., Liedtke, C., Tutt, A. and von Minckwitz, G., 2017. Molecular alterations in triple-negative breast cancer-the road to new treatment strategies. *The Lancet,* 389(10087) pp. 2430 - 2442.

Desai, N.V. et al., 2018. Intrinsic molecular subtypes of breast cancers categorized as HER2-positive using an alternative chromosome 17 probe assay. *Breast Cancer Research,* 20(1) pp. 75.

Di Leva, G. and Croce, C.M., 2013a. miRNA profiling of cancer. *Current Opinion in Genetics and Development,* 23(1) pp. 3 - 11.

Di Leva, G. and Croce, C.M. 2013b. miRNAs in Cancer Initiation and Progression. pp. 57-62.

Di Resta, C. and Ferrari, M., 2018. Next generation sequencing: from research area to clinicalpractice. *International Federation of Clinical Chemistry and Laboratory Medicine,* 29(3) pp. 215 - 220.

Ding, S. et al., 2015. Redox regulation in cancer stem cells. *Oxidative Medicine and Cellular Longevity,* 2015 pp. e750798.

Dobbin, Z.C. and Landen, C.N., 2013. Isolation and characterization of potential cancer stem cells from solid human tumors--potential applications. *Current protocols in pharmacology,* 63 pp. Unit-14.28.

Dobson, J.R. et al., 2014. hsa-mir-30c promotes the invasive phenotype of metastatic breast cancer cells by targeting NOV/CCN3. *Cancer Cell International,* 14(1) pp. 73 - 87.

Douville, J., Beaulieu, R. and Balicki, D., 2008. ALDH1 as a functional marker of cancer stem and progenitor cells. *Stem Cells and Development,* 18(1) pp. 17 - 26.

Drahos, J. et al., 2016. MicroRNA profiles of Barrett's esophagus and esophageal adenocarcinoma: differences in glandular non-native epithelium. *Cancer Epidemiology, Biomarkers and Prevention,* 25(3) pp. 429 - 437.

Drexler, H. G. and Uphoff, C. C., 2002. Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention. *Cytotechnology,* 39(2) pp. 75 – 90.

Du, H.-Y. and Liu, B., 2018. MiR-1271 as a tumor suppressor in breast cancer proliferation and progression via targeting SPIN1. *European Review for Medical and Pharmacological Sciences.,* 22(9) pp. 2697 - 2706.

Duan, J.-j. et al., 2013. Strategies for isolating and enriching cancer stem cells: well begun is half done. *Stem Cells and Development,* 22(16) pp. 2221 - 2239.

Edmondson, R., Broglie, J.J., Adcock, A.F. and Yang, L., 2014. Three-dimensional cell culture systems and their applications in drug discovery and cell-based biosensors. *Assay and Drug Development Technologies,* 12(4) pp. 207 - 218.

Ehrlich, M., 2009. DNA hypomethylation in cancer cells. *Epigenomics,* 1(2) pp. 239 - 259.

Eichner, L.J. et al., 2010. miR-378* mediates metabolic shift in breast cancer cells via the PGC-1β/ERRγ transcriptional pathway. *Cell Metabolism,* 12(4) pp. 352 - 361.

Ekert, J.E. et al., 2013. Three-dimensional lung tumor microenvironment modulates therapeutic compound responsiveness *in vitro* implication for drug development. *PLoS ONE,* 9(3) pp. e92248.

El-Khoury, V. et al., 2016. Assessing cellular and circulating miRNA recovery: the impact of the RNA isolation method and the quantity of input material. *Scientific Reports,* 6 pp. e19529.

El Helou, R. et al., 2014. Brief reports: a distinct DNA methylation signature defines breast cancer stem cells and predicts cancer outcome. *Stem Cells,* 32(11) pp. 3031 - 3036.

Elgamal, O.A. et al., 2013. Tumor suppressive function of mir-205 in breast cancer is linked to HMGB3 regulation. *PLoS ONE,* 8(10) pp. e76402.

Eliyatkın, N. et al., 2015. Molecular classification of breast carcinoma: from traditional, old-fashioned way to a new age, and a new way. *The Journal of Breast Health,* 11(2) pp. 59 - 66.

Ell, B. et al., 2014. The microRNA-23b/27b/24 cluster promotes breast cancer lung metastasis by targeting metastasis-suppressive gene prosaposin. *The Journal of Biological Chemistry,* 289(32) pp. 21888 - 21895.

Endrullat, C., Glokler, J., Franke, P. and Frohme, M., 2016. Standardization and quality management in next-generation sequencing. *Applied and Translational Genomics,* 10 pp. 2 - 9.

Esteller, M., 2011. Non-coding RNAs in human disease. *Nature Reviews Genetics,* 12 pp. 861 - 874.

Ewing, B., Hillier, L., Wendl, M.C. and Green, P., 1998. Base-calling of automated sequencer traces using Phred.I. accuracy assessment. *Genome Research,* 8(3) pp. 175 - 185.

Fan, X. et al., 2017. MicroRNAs, a subpopulation of regulators, are involved in breast cancer progression through regulating breast cancer stem cells. *Oncology Letters,* 14(5) pp. 5069 - 5076.

Fan, X. et al., 2013. Up-regulated microRNA-143 in cancer stem cells differentiation promotes prostate cancer cells metastasis by modulating FNDC3B expression. *BioMed Central Cancer,* 13 pp. e61.

Fan, X. et al., 2010. Effective enrichment of prostate cancer stem cells from spheres in a suspension culture system. *Urologic oncology,* 30(3) pp. 314 - 318.

Fanali, C. et al., 2014. Cancer stem cells in colorectal cancer from pathogenesis to therapy: controversies and perspectives. *World Journal of Gastroenterology,* 20(4) pp. 923 - 942.

Farahani, E. et al., 2014. Cell adhesion molecules and their relation to (cancer) cell stemness. *Carcinogenesis,* 35(4) pp. 747 - 759.

Feliciano, A. et al., 2013. miR-125b acts as a tumor suppressor in breast tumorigenesis via its novel direct targets ENPEP, CK2-α, CCNJ, and MEGF9. *PLoS ONE,* 8(10) pp. e76247.

Feng, X., Wang, Z., Fillmore, R. and Xi, Y., 2014. MiR-200, a new star miRNA in human cancer. *Cancer Letters,* 344(2) pp. 166 - 173.

Feng, Y. et al., 2018. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes and Diseases,* 5(2) pp. 77 - 106.

Ferlay, J. et al., 2018. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *European Journal of Cancer,* 103 pp. 356 - 387.

Fillmore, C. and Kuperwasser, C., 2007. Human breast cancer stem cell markers CD44 and CD24: enriching for cells with functional properties in mice or in man? *Breast Cancer Research,* 9(3) pp. 303 - 305.

Finn, R.S. et al., 2016. Palbociclib and letrozole in advanced breast cancer. *New England Journal of Medicine,* 375(20) pp. 1925 - 1936.

Fiori, M.E. et al., 2019. Cancer-associated fibroblasts as abettors of tumor progression at the crossroads of EMT and therapy resistance. *Molecular Cancer,* 18(1) pp. e70.

Fiorillo, M., Sotgia, F. and Lisanti, M.P., 2019. 'Energetic' cancer stem cells (e-CSCs): a new hyper-metabolic and proliferative tumor cell phenotype, driven by mitochondrial energy. *Frontiers in Oncology,* 8 pp. e00677.

Fix LN et al., 2010. MicroRNA expression profile of MCF-7 human breast cancer cells and the effect of green tea polyphenon-60. *Cancer Genomics Proteomics,* 7(5) pp. 261 - 277.

Florl, A.R. et al., 2004. Coordinate hypermethylation at specific genes in prostate carcinoma precedes LINE-1 hypomethylation. *British Journal of Cancer,* 91(5) pp. 985 - 994.

Frixa, T., Donzelli, S. and Blandino, G., 2015. Oncogenic microRNAs: key players in malignant transformation. *Cancers,* 7(4) pp. 2466 - 2485.

Galli, R. 2019. The neurosphere assay (NSA) applied to neural stem cells (NSCs) and cancer stem cells (CSCs). *Target Identification and Validation in Drug Discovery: Methods and Protocols*. New York, NY, Springer New York. pp. 139 - 149.

Gallo, A. et al., 2018. On the prospect of serum exosomal miRNA profiling and protein biomarkers for the diagnosis of ascending aortic dilatation in patients with bicuspid and tricuspid aortic valve. *International Journal of Cardiology,* 273 pp. 230 - 236.

Gan, R. et al., 2014. Downregulation of miR-221/222 enhances sensitivity of breast cancer cells to tamoxifen through upregulation of TIMP3. *Cancer Gene Therapy,* 21(7) pp. 290 - 296.

Gao, H.-Y. et al., 2018. Screening of prognostic risk microRNAs for acute myeloid leukemia. *Hematology,* 23(10) pp. 747 - 755.

Gao, W. et al., 2018. Development of a novel and economical agar-based non-adherent three-dimensional culture method for enrichment of cancer stem-like cells. *Stem Cell Research and Therapy,* 9(1) pp. 243.

Garcia-Etienne, C.A. et al., 2012. Mastectomy trends for early-stage breast cancer: A report from the EUSOMA multi-institutional European database. *European Journal of Cancer,* 48(13) pp. 1947 - 1956.

Garg, M., 2015. Emerging role of microRNAs in cancer stem cells: Implications in cancer therapy. *World journal of stem cells,* 7(8) pp. 1078 - 1089.

Gasparini, P. et al., 2014. microRNA expression profiling identifies a four microRNA signature as a novel diagnostic and prognostic biomarker in triple negative breast cancers. *Oncotarget,* 5(5) pp. 1174 - 1184.

Gedye, C. et al., 2016. Mycoplasma infection alters cancer stem cell properties *in vitro*. *Stem Cell Reviews and Reports,* 12(1) pp. 156 - 161.

Gencoglu, M.F. et al., 2018. Comparative study of multicellular tumor spheroid formation methods and implications for drug screening. *American Chemical Society Biomaterials Science and Engineering,* 4(2) pp. 410 - 420.

Geng, S. et al., 2013. Cancer stem-like cells enriched with CD29 and CD44 markers exhibit molecular characteristics with epithelial-mesenchymal transition in squamous cell carcinoma. *Archives of Dermatological Research,* 305(1) pp. 35 - 47.

Geurts, Y.M. et al., 2017. Patterns and predictors of first and subsequent recurrence in women with early breast cancer. *Breast Cancer Research and Treatment,* 165(3) pp. 709 - 720.

Giacomelli, A.O. et al., 2018. Mutational processes shape the landscape of TP53 mutations in human cancer. *Nature Genetics,* 50(10) pp. 1381 - 1387.

Giglia-Mari, G., Zotter, A. and Vermeulen, W., 2011. DNA damage response. *Cold Spring Harbor perspectives in biology,* 3(1) pp. a000745.

Gilmore, A.P., 2005. Anoikis. *Cell Death And Differentiation,* 12 pp. 1473 - 1477.

Gilmore, E., McCabe, N., Kennedy, R.D. and Parkes, E.E., 2019. DNA repair deficiency in breast cancer: opportunities for immunotherapy. *Journal of Oncology,* 2019 pp. e14.

Ginestier, C. et al., 2007. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell,* 1(5) pp. 555 - 567.

Git, A. et al., 2009. Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA,* 16(5) pp. 991 - 1006.

Giza, D.E., Vasilescu, C. and Calin, G.A., 2014. Key principles of miRNA involvement in human diseases. *Discoveries (Craiova, Romania),* 2(4) pp. e34.

Gomes, A. et al., 2016. Oxygen partial pressure is a rate-limiting parameter for cell proliferation in 3D spheroids grown in physioxic culture condition. *PLoS ONE,* 11(8) pp. e0161239.

Grimshaw, M. et al., 2008. Mammosphere culture of metastatic breast cancer cells enriches for tumorigenic breast cancer cells. *Breast Cancer Research,* 10(3) pp. R52.

Grosso, G. et al., 2017. Possible role of diet in cancer: systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. *Nutrition Reviews,* 75(6) pp. 405 - 419.

Gu, G., Yuan, J., Wills, M. and Kasper, S., 2007. Prostate cancer cells with stem cell characteristics reconstitute the original human tumor in vivo. *Cancer Research,* 67(10) pp. 4807 - 4815.

Gu, Y. et al., 2014. MicroRNA-493 suppresses tumor growth, invasion and metastasis of lung cancer by regulating E2F1. *PLoS ONE,* 9(8) pp. e102602.

Gu, Y. et al., 2018. miR-3074-5p promotes the apoptosis but Inhibits the invasiveness of human extravillous trophoblast-derived HTR8/SVneo cells *in vitro*. *Reproductive Sciences,* 25(5) pp. 690 - 699.

Guénard, F. et al., 2017. Genome-wide association study of dietary pattern scores. *Nutrients,* 9(7) pp. E649.

Guo, C. et al., 2012. Epcam, CD44, and CD49f distinguish sphere-forming human prostate basal cells from a subpopulation with predominant tubule initiation capability. *PLoS ONE,* 7(4) pp. e34219.

Guo, L. et al., 2016. miRNA-411 acts as a potential tumor suppressor miRNA via the downregulation of specificity protein 1 in breast cancer. *Molecular Medicine Reports,* 14(4) pp. 2975 - 2982.

Guo, X. et al., 2019. Enrichment of cancer stem cells by agarose multi-well dishes and 3D spheroid culture. *Cell and Tissue Research,* 375(2) pp. 397 - 408.

Han, M. et al., 2012. MiR-21 regulates epithelial-mesenchymal transition phenotype and hypoxia-inducible factor-1α expression in third-sphere forming breast cancer stem cell-like cells. *Cancer Science,* 103(6) pp. 1058 - 1064.

Han, Y. and Li, H., 2018. miRNAs as biomarkers and for the early detection of non-small cell lung cancer (NSCLC). *Journal of Thoracic Disease,* 10(5) pp. 3119 - 3131.

Hanahan, D. and Weinberg, R.A., 2000. The hallmarks of cancer. *Cell,* 100(1) pp. 57 - 70.

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

Hanna, J., Hossain, G.S. and Kocerha, J., 2019. The potential for microRNA therapeutics and clinical research. *Frontiers in Genetics,* 10 pp. e478.

Hao, G.-j. et al., 2017. Attenuation of deregulated miR-369-3p expression sensitizes non-small cell lung cancer cells to cisplatin via modulation of the nucleotide sugar transporter SLC35F5. *Biochemical and Biophysical Research Communications,* 488(3) pp. 501 - 508.

Hao, J. et al., 2014. MicroRNA control of epithelial-mesenchymal transition in cancer stem cells. *International Journal of Cancer*, 135(5) pp. 1019 - 1027.

Hassanpour, S.H. and Dehghani, M., 2017. Review of cancer from perspective of molecular. *Journal of Cancer Research and Practice*, 4(4) pp. 127 - 129.

Hawkes, N., 2019. Cancer survival data emphasise importance of early diagnosis. *British Medical Journal*, 364 pp. e1408.

Hayes, J., Peruzzi, P.P. and Lawler, S., 2014. MicroRNAs in cancer: biomarkers, functions and therapy. *Trends in Molecular Medicine*, 20(8) pp. 460 - 469.

He, X. et al., 2016. MiR-381 functions as a tumor suppressor in colorectal cancer by targeting Twist1. *OncoTargets and therapy*, 9 pp. 1231 - 1239.

He, Y.-J. et al., 2013. miR-342 is associated with estrogen receptor-Î± expression and response to tamoxifen in breast cancer. *Experimental and Therapeutic Medicine*, 5(3) pp. 813 - 818.

Head, S.R. et al., 2014. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*, 56(2) pp. 61 - 92.

Herheliuk, T., et al., 2019. Investigation of multicellular tumour spheroids enriched for a cancer stem cell phenotype. *Stem Cell Investigation*, 6 p. e21.

Hibino, S. et al., 2014. Inhibitors of enhancer of zeste homolog 2 (EZH2) activate tumor-suppressor microRNAs in human cancer cells. *Oncogenesis*, 3(5) pp. e104.

Hilton, H.N., Clarke, C.L. and Graham, J.D., 2018. Estrogen and progesterone signalling in the normal breast and its implications for cancer development. *Molecular and Cellular Endocrinology*, 466 pp. 2 - 14.

Hirata, A. et al., 2019. Heterogeneity of colorectal cancer stem cells. *Cancer Prevention Research*, 18 pp. e0482.

Hirschhaeuser, F. et al., 2010. Multicellular tumor spheroids: An underestimated tool is catching up again. *Journal of Biotechnology*, 148(1) pp. 3 - 15.

Ho, C.S., Noor, S.M. and Nagoor, N.H., 2018. MiR-378 and miR-1827 regulate tumor invasion, migration and angiogenesis in human lung adenocarcinoma by targeting RBX1 and CRKL, respectively. *Journal of Cancer*, 9(2) pp. 331 - 345.

Ho, J.-Y. et al., 2016. Estrogen enhances the cell viability and motility of breast cancer cells through the ERα-ΔNp63-Integrin β4 signaling pathway. *PLoS ONE*, 11(2) pp. e0148301.

Ho, W.Y. et al., 2012. Development of multicellular tumor spheroid (MCTS) culture from breast cancer cell and a high throughput screening method using the MTT assay. *PLoS ONE,* 7(9) pp. e44640.

Holliday, D. and Speirs, V., 2011. Choosing the right cell line for breast cancer research. *Breast Cancer Research,* 13(4) pp. 215 - 222.

Honoki, K. et al., 2010. Possible involvement of stem-like populations with elevated ALDH1 in sarcomas for chemotherapeutic drug resistance. *Oncology Reports,* 24(2) pp. 501 - 505.

Hoppe, R. et al., 2013. Increased expression of miR-126 and miR-10a predict prolonged relapse-free time of primary oestrogen receptor-positive breast cancer following tamoxifen treatment. *European Journal of Cancer,* 49(17) pp. 3598 - 3608.

Horimoto, Y. et al., 2016. Combination of cancer stem cell markers CD44 and CD24 is superior to ALDH1 as a prognostic indicator in breast cancer patients with distant metastases. *PLoS ONE,* 11(10) pp. e0165253.

Hou, G.-X. et al., 2018. Elimination of stem-like cancer cell side-population by auranofin through modulation of ROS and glycolysis. *Cell Death and Disease,* 9(2) pp. 89.

Hu, J. et al., 2019. Cancer stem cell self-renewal as a therapeutic target in human oral cancer. *Oncogene,* 38(27) pp. 5440 - 5456.

Hu S.-H. et al., 2016. miR-760 mediates chemoresistance through inhibition of epithelial mesenchymal transition in breast cancer cells. *European Review for Medical and Pharmacological Sciences,* 20(23) pp. 5002 - 5008.

Hu, X. et al., 2012. Induction of cancer cell stemness by chemotherapy. *Cell Cycle,* 11(14) pp. 2691 - 2698.

Hu Y., Lan W. and D., M. 2017. *Next-generation sequencing for microRNA expression profile*. New York: Humana Press.

Hua, Z. et al., 2006. MiRNA-directed regulation of VEGF and other angiogenic factors under hypoxia. *PLoS ONE,* 1(1) pp. e116.

Huang, R. and Rofstad, E.K., 2017. Cancer stem cells (CSCs), cervical CSCs and targeted therapies. *Oncotarget,* 8(21) pp. 35351 - 35367.

Huang, S. et al., 2012. miR-143 and miR-145 inhibit stem cell characteristics of PC-3 prostate cancer cells. *Oncology Reports,* 28(5) pp. 1831 -1837.

Hutchinson, L., 2010. Challenges, controversies, breakthroughs. *Nature Reviews Clinical Oncology,* 7 pp. 669.

Hynds, R.E., Vladimirou, E. and Janes, S.M., 2018. The secret lives of cancer cell lines. *Disease Models and Mechanisms.,* 11(11) pp. dmm037366.

Ikeda, Y. et al., 2012. MicroRNAs associated with mitogen-activated protein kinase in human pancreatic cancer. *Molecular Cancer Research,* 10(2) pp. 259 - 269.

Imamura, Y. et al., 2015 Comparison of 2D- and 3D-culture models as drug-testing platforms in breast cancer. *Oncology Reports,* 33 pp. 1837 - 1843.

Ingangi, V. et al., 2019. Role of microenvironment on the fate of disseminating cancer stem cells. *Frontiers in Oncology,* 9 pp. e82.

Iorio, M.V. et al., 2005. MicroRNA gene expression deregulation in human breast cancer. *Cancer Research,* 65(16) pp. 7065 - 7070.

Ishiguro, T. et al., 2017. Tumor-derived spheroids: Relevance to cancer stem cells and clinical applications. *Cancer Science,* 108(3) pp. 283 - 289.

Ishiwata, T., Matsuda, Y. and Yoshimura, H., 2018. Pancreatic cancer stem cells: features and detection methods. *Pathology and Oncology Research,* 24 pp. 797.

Islami, F., Torre, L.A. and Jemal, A., 2015. Global trends of lung cancer mortality and smoking prevalence. *Translational Lung Cancer Research,* 4(4) pp. 327 - 338.

Isobe, T. et al., 2014. miR-142 regulates the tumorigenicity of human breast cancer stem cells through the canonical WNT signaling pathway. *eLife,* 3 pp. e01977.

Ivascu, A. and Kubbies, M., 2007. Diversity of cell-mediated adhesions in breast cancer spheroids. *International Journal of Oncology,* 31(6) pp. 1403 - 1413.

Jaggupilli, A. and Elkord, E., 2012. Significance of CD44 and CD24 as cancer stem cell markers: an enduring ambiguity. *Clinical and Developmental Immunology,* 2012 pp. 1-11.

Javed, A. and Lteif, A., 2013. Development of the human breast. *Seminars in Plastic Surgery,* 27(1) pp. 5 -12.

Jiang, F. et al., 2014. The repressive effect of miR-148a on TGF beta-SMADs signal pathway is involved in the glabridin-induced inhibition of the cancer stem cells-like properties in hepatocellular carcinoma cells. *PLoS ONE,* 9(5) pp. e96698.

Jiang, Y.X. et al., 2017. The promotion of the transformation of quiescent gastric cancer stem cells by IL-17 and the underlying mechanisms. *Oncogene,* 36(9) pp. 1256 - 1264.

Jiang, Z.-B. et al., 2019. miR-365 regulates liver cancer stem cells via RAC1 pathway. *Molecular Carcinogenesis,* 58(1) pp. 55 - 65.

Jin, L. et al., 2013. Prooncogenic factors miR-23b and miR-27b are regulated by Her2/Neu, EGF, and TNF in breast cancer. *Cancer Research,* 73(9) pp. 2884 - 2896.

Jin, X. and Mu, P., 2015. Targeting breast cancer metastasis. *Breast cancer :Basic and Clinical Research,* 9(Suppl 1) pp. 23 - 34.

Jing, H. et al., 2016. Suppression of Spry4 enhances cancer stem cell properties of human MDA-MB-231 breast carcinoma cells. *Cancer Cell International,* 17(53) pp. 28507453.

Josson, S. et al., 2014. miR-409-3p/-5p promotes tumorigenesis, epithelial to mesenchymal transition and bone metastasis of human prostate cancer. *Clinical Cancer Research,* 20(17) pp. 4636 - 4646.

Ju, J., Zhu, A.-J. and Yuan, P., 2018. Progress in targeted therapy for breast cancer. *Chronic Diseases and Translational Medicine,* 4(3) pp. 164 - 175.

Jung, M. et al., 2010. Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. *Clinical Chemistry,* 56(6) pp. 998 - 1006.

Kademani, D. et al., 2009. Angiogenesis and CD34 expression as a predictor of recurrence in oral squamous cell carcinoma. *Journal of Oral and Maxillofacial Surgery,* 67(9) pp. 1800 - 1805.

Kalia, M., 2015. Biomarkers for personalized oncology: recent advances and future challenges. *Metabolism - Clinical and Experimental,* 64(3) pp. S16 - S21.

Kang, L. et al., 2015. MicroRNA-34a suppresses the breast cancer stem cell-like characteristics by downregulating Notch1 pathway. *Cancer science,* 106(6) pp. 700 - 708.

Kang, L. et al., 2015. Aberrant allele-switch imprinting of a novel IGF1R intragenic antisense non-coding RNA in breast cancers. *European Journal of Cancer,* 51(2) pp. 260 - 270.

Kao, J. et al., 2009. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS ONE,* 4(7) pp. e6146.

Karlsson, P. et al., 2017. Timing of radiation therapy and chemotherapy after breast-conserving surgery for node-positive breast cancer: long-term results from international breast cancer study group trials VI and VII. *International Journal of Radiation Oncology, Biology, Physics,* 96(2) pp. 273 - 279.

Katoh, M., 2017. Canonical and non-canonical WNT signaling in cancer stem cells and their niches: Cellular heterogeneity, omics reprogramming, targeted therapy and tumor plasticity (Review). *International Journal of Oncology,* 51(5) pp. 1357 - 1369.

Ke, X. et al., 2017. MiR-410 induces stemness by inhibiting Gsk3β but upregulating β-catenin in non-small cells lung cancer. *Oncotarget,* 8(7) pp. 11356 - 11371.

Kelley, M.R., Logsdon, D. and Fishel, M.L., 2014. Targeting DNA repair pathways for cancer treatment: what's new? *Future Oncology,* 10(7) pp. 1215 - 1237.

Kim, G. et al., 2016. Hsa-miR-1246 and hsa-miR-1290 are associated with stemness and invasiveness of non-small cell lung cancer. *Lung Cancer,* 91 pp. 15 - 22.

Kim, H.S., Pearson, A.T. and Nör, J.E., 2016. Isolation and characterization of cancer stem cells from primary head and neck squamous cell carcinoma tumors. *Methods in Molecular Biology,* 1395 pp. 241 - 249.

Kim JB, Stein R and MJ, O.H., 2004. Three-dimensional *in vitro* tissue culture models of breast cancer-a review. *Breast Cancer Research and Treatment,* 85 pp. 281 - 291.

Kim, M.K. et al., 2019. Mycoplasma infection promotes tumor progression via interaction of the mycoplasmal protein p37 and epithelial cell adhesion molecule in hepatocellular carcinoma. *Cancer Letters,* 454 pp. 44 - 52.

Kim, S.-J. et al., 2012. MicroRNA let-7a suppresses breast cancer cell migration and invasion through downregulation of C-C chemokine receptor type 7. *Breast Cancer Research,* 14(1) pp. R14.

Kim, T. et al., 2011. p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. *The Journal of Experimental Medicine,* 208(5) pp. 875 - 883.

Kim, W.-T. and Ryu, C.J., 2017. Cancer stem cell surface markers on normal stem cells. *Journal of Biochemistry and Molecular Biology,* 50(6) pp. 285 - 298.

Klevebring, D. et al., 2014. Sequencing of breast cancer stem cell populations indicates a dynamic conversion between differentiation states in vivo. *Breast Cancer Research,* 16(4) pp. R72.

Kong, Y. et al., 2018. Breast cancer stem cell markers CD44 and ALDH1A1 in serum: distribution and prognostic value in patients with primary breast cancer. *Journal of Cancer,* 9(20) pp. 3728 - 3735.

Kopan, R., 2012. Notch signaling. *Cold Spring Harbor perspectives in biology,* 4(10) pp. a011213.

Korch, C. et al., 2018. Authentication of M14 melanoma cell line proves misidentification of MDA-MB-435 breast cancer cell line. *International Journal of Cancer,* 142(3) pp. 561 - 572.

Koshiol, J. et al., 2010. Strengths and limitations of laboratory procedures for microRNA detection. *Cancer Epidemiology, Biomarkers and Prevention,* 19(4) pp. 907 - 911.

Koury, J., Zhong, L. and Hao, J., 2017. Targeting signaling pathways in cancer stem cells for cancer treatment. *Stem Cells International,* 2017 pp. 2925869.

Kovats, S., 2015. Estrogen receptors regulate innate immune cells and signaling pathways. *Cellular Immunology,* 294(2) pp. 63 - 69.

Krishnan, K. et al., 2013. miR-139-5p is a regulator of metastatic pathways in breast cancer. *RNA (New York, N.Y.),* 19(12) pp. 1767 - 1780.

Kristensen, V. 2015. *Integrated molecular analysis of miRNAs in breast cancer with clinical outcome* [*online*] [Accessed 18 October 2019.

Kruyt, F.A.E. and Schuringa, J.J., 2010. Apoptosis and cancer stem cells: Implications for apoptosis targeted therapy. *Biochemical Pharmacology,* 80(4) pp. 423 - 430.

Kumar, B. et al., 2018. Normal breast-derived epithelial cells with luminal and intrinsic subtype-enriched gene expression document interindividual differences in their differentiation cascade. *Cancer Research,* 78(17) pp. 5107.

La Porta, C.A., 2012. Thoughts about cancer stem cells in solid tumors. *World Journal of Stem Cells,* 4(3) pp. 17 - 20.

Lacroix, M. and Leclercq, G., 2004. Relevance of breast cancer cell lines as models for breast tumours: an update. *Breast Cancer Research and Treatment,* 83(3) pp. 249 - 289.

Langhans, S.A., 2018. Three-dimensional *in vitro* cell culture models in drug discovery and drug repositioning. *Frontiers in Pharmacology,* 9(6) pp. 29410625.

Lapidot, T. et al., 1994. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature,* 367(6464) pp. 645-648.

Laplante, M. and Sabatini, D.M., 2009. mTOR signaling at a glance. *Journal of Cell Science,* 122(20) pp. 3589.

Lawless, N. et al., 2013. Next generation sequencing reveals the expression of a unique miRNA profile in response to a gram-positive bacterial infection. *PLoS ONE,* 8(3) pp. e57543.

Lawson, D.A. et al., 2018. Tumour heterogeneity and metastasis at single-cell resolution. *Nature cell biology,* 20(12) pp. 1349 - 1360.

Lechman, E.R. et al., 2016. miR-126 regulates distinct self-renewal outcomes in normal and malignant hematopoietic stem cells. *Cancer cell,* 29(2) pp. 214 - 228.

Ledford, H., 2016. US cancer institute to overhaul tumour cell lines. *Nature,* 530(7591) pp. 391.

Lee, H. et al., 2016. Genome-wide identification of target genes for miR-204 and miR-211 identifies their proliferation stimulatory role in breast cancer cells. *Scientific Reports,* 6 pp. e25287.

Lee, I.-H. et al., 2014. A detailed analysis of next generation sequencing reads of microRNA expression in Barrett's esophagus: absolute versus relative quantification. *BMC Research Notes,* 7 pp. e212.

Lee, J.M. et al., 2018. Generation of uniform-sized multicellular tumor spheroids using hydrogel microwells for advanced drug screening. *Scientific Reports,* 8(1) pp. 17145.

Lee, J.W. et al., 2018. Isolation of spheroid-forming single cells from gastric cancer cell lines: enrichment of cancer stem-like cells. *Biotechniques,* 65(4) pp. 197 - 203.

Levy, S.E. and Myers, R.M., 2016. Advancements in next-generation sequencing. *Annual Review of Genomics and Human Genetics,* 17(1) pp. 95 - 115.

Li, B. et al., 2016. miR-221/222 enhance the tumorigenicity of human breast cancer stem cells via modulation of PTEN/Akt pathway. *Biomedicine and Pharmacotherapy,* 79 pp. 93 - 101.

Li, H., Zhang, S., Wang, Q. and Zhu, R., 2016. Clinical value of mammography in diagnosis and identification of breast mass. *Pakistan Journal of Medical Sciences,* 32(4) pp. 1020 - 1025.

Li, J.-H. et al., 2018. Diagnostic and prognostic value of microRNA-628 for cancers. *Journal of Cancer,* 9(9) pp. 1623 - 1634.

Li, L. et al., 2013. MiR-34a inhibits proliferation and migration of breast cancer through down-regulation of Bcl-2 and SIRT1. *Clinical and Experimental Medicine,* 13(2) pp. 109 -117.

Li, M. et al., 2018. MiR-628-5p decreases the tumorigenicity of epithelial ovarian cancer cells by targeting at FGFR2. *Biochemical and Biophysical Research Communications,* 495(2) pp. 2085 - 2091.

Li, T. et al., 2010. ALDH1A1 is a marker for malignant prostate stem cells and predictor of prostate cancer patients's outcome. *Laboratory Investigation,* 90(2) pp. 234 - 244.

Li, W. et al., 2014. Decreased expression of miR-204 is associated with poor prognosis in patients with breast cancer. *International Journal of Clinical and Experimental Pathology,* 7(6) pp. 3287 - 3292.

Li, W. et al., 2017. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. *Scientific Reports,* 7(1) pp. 13856 - 13856.

Li, X.J., Ren, Z.J. and Tang, J.H., 2014. MicroRNA-34a: a potential therapeutic target in human cancer. *Cell Death and Disease,* 5 pp. e1327.

Li, X.J., Ren, Z.J., Tang, J.H. and Yu, Q., 2017. Exosomal microRNA miR-1246 promotes cell proliferation,invasion and drug resistance by targeting CCNG2 in breast Cancer. *Cellular Physiology and Biochemistry,* 44(5) pp. 1741 - 1748.

Li, Y., Deng, X., Zeng, X. and Peng, X., 2016. The role of mir-148a in cancer. *Journal of Cancer,* 7(10) pp. 1233 - 1241.

Li, Y. and Kowdley, K.V., 2012. MicroRNAs in common human diseases. *Genomics, proteomics & bioinformatics,* 10(5) pp. 246 - 253.

Li, Y. et al., 2010. Sulforaphane, a dietary component of broccoli/broccoli sprouts, inhibits breast cancer stem cells. *Clinical Cancer Research,* 16(9) pp. 2580 - 2590.

Liang, D.-S. et al., 2018. Vitamin E-based redox-sensitive salinomycin prodrug-nanosystem with paclitaxel loaded for cancer targeted and combined chemotherapy. *Colloids and Surfaces B: Biointerfaces,* 172 pp. 506 - 516.

Liao, J. et al., 2014. Ovarian cancer spheroid cells with ctem cell-like properties contribute to tumor generation, metastasis and chemotherapy resistance through hypoxia-resistant metabolism. *PLoS ONE,* 9(1) pp. e84941.

Lim, S.-O. et al., 2016. EGFR signaling enhances aerobic glycolysis in triple-negative breast cancer cells to promote tumor growth and immune escape. *Cancer Research,* 76(5) pp. 1284 - 1296.

Lin, C. et al., 2018. MicroRNA 628 suppresses migration and invasion of breast cancer stem cells through targeting SOS1. *OncoTargets and therapy,* 11 pp. 5419 - 5428.

Lin, S.-P. et al., 2012. Survival of cancer stem cells under hypoxia and serum depletion via decrease in PP2A activity and activation of p38-MAPKAPK2-Hsp27. *PLoS ONE,* 7(11) pp. e49605.

Lin, S. and Gregory, R.I., 2015. MicroRNA biogenesis pathways in cancer. *Nature Reviews. Cancer,* 15(6) pp. 321- 333.

Lin, X. et al., 2017. Nanoparticle delivery of miR-34a eradicates long-term-cultured breast cancer stem cells via targeting C22ORF28 directly. *Theranostics,* 7(19) pp. 4805 - 4824.

Lin, Y. et al., 2015. MicroRNA-33b inhibits breast cancer metastasis by targeting HMGA2, SALL4 and Twist1. *Scientific Reports,* 5 pp. e9995.

Lipovka, Y. and Konhilas, J.P., 2016. The complex nature of oestrogen signalling in breast cancer: enemy or ally? *Bioscience Reports,* 36(3) pp. e00352.

Liu, C. et al., 2011. The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nature medicine,* 17(2) pp. 211 - 215.

Liu, F. et al., 2016. Biomarkers for EMT and MET in breast cancer: An update. *Oncology letters,* 12(6) pp. 4869 - 4876.

Liu, H. et al., 2018. Identifying and targeting sporadic oncogenic genetic aberrations in mouse models of triple-negative breast cancer. *Cancer Discovery,* 8(3) pp. 354.

Liu, J. et al., 2014. MicroRNA expression profile of gastric cancer stem cells in the MKN-45 cancer cell line. *Acta Biochimica et Biophysica Sinica,* 46(2) pp. 92 - 99.

Liu, J. et al., 2013. Spheroid body-forming cells in the human gastric cancer cell line MKN-45 possess cancer stem cell properties. *International Journal of Oncology,* 42(2) pp. 453 -459.

Liu, J. et al., 2013. Significantly increased expression of OCT4 and ABCG2 in spheroid body-forming cells of the human gastric cancer MKN-45 cell line. *Oncology letters,* 6(4) pp. 891- 896.

Liu, J. et al., 2013. Lung cancer tumorigenicity and drug resistance are maintained through ALDH(hi)CD44(hi) tumor initiating cells. *Oncotarget,* 4(10) pp. 1698 -1711.

Liu, L. et al., 2018. Chemotherapy induces breast cancer stemness in association with dysregulated monocytosis. *Clinical Cancer Research,* 24(10) pp. 2370.

Liu, P. 2018. *MicroRNA expression analysis: Next-Generation Sequencing*. New York: Humana Press.

Liu, S. et al., 2014. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports,* 2(1) pp. 78 - 91.

Liu, X.-Y. et al., 2015. Induction of autophagy and apoptosis by miR-148a through the sonic hedgehog signaling pathway in hepatic stellate cells. *American journal of cancer research,* 5(9) pp. 2569 - 2589.

Liu, X. et al., 2016. Targeting ALDH1A1 by disulfiram/copper complex inhibits non-small cell lung cancer recurrence driven by ALDH-positive cancer stem cells. *Oncotarget,* 7(36) pp. 58516 - 58530.

Liu, Y. et al., 2015. TP53 loss creates therapeutic vulnerability in colorectal cancer. *Nature,* 520(7549) pp. 697 - 701.

Liu, Y., Zhou, J. and White, K.P., 2014. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics,* 30(3) pp. 301 - 304.

Lo, S.S. et al., 2012. Overexpression of miR-370 and downregulation of its novel target TGFβ-RII contribute to the progression of gastric carcinoma. *Oncogene,* 31(2) pp. 226 - 237.

Loef, M. and Walach, H., 2012. The combined effects of healthy lifestyle behaviors on all cause mortality: A systematic review and meta-analysis. *Preventive Medicine,* 55(3) pp. 163 -170.

Lopez, J. et al., 2012. Cancer-initiating cells derived from established cervical cell lines exhibit stem-cell markers and increased radioresistance. *BioMed Central Cancer,* 12(1) pp. 48.

Love C. and S., D. 2013. *MicroRNA expression profiling using microarrays*. Totowa, NJ: Humana Press.

Lowery, A.J. et al., 2009. MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer. *Breast Cancer Research : BCR,* 11(3) pp. R27.

Lü, L. et al., 2017. MicroRNAs in the prognosis of triple-negative breast cancer: A systematic review and meta-analysis. *Medicine,* 96(22) pp. e7085.

Lu, Y. et al., 2014. MiR-200a inhibits epithelial-mesenchymal transition of pancreatic cancer stem cell. *BioMed Central Cancer,* 14 pp. e85.

Luca, A.C. et al., 2013. Impact of the 3D microenvironment on phenotype, gene expression, and EGFR inhibition of colorectal cancer cell lines. *PLoS ONE,* 8(3) pp. e59689.

Ludwig, K., Tse, E. and Wang, J., 2013. Colon cancer cells adopt an invasive phenotype without mesenchymal transition in 3-D but not 2-D culture upon combined stimulation with EGF and crypt growth factors. *BioMed Central Cancer,* 13(1) pp. 221 - 233.

Lum, M. 2019. Breast cancer reality check: We’re dying more than we should. *The Star*. Kuala Lumpur.

Lund, A.H., 2009. miR-10 in development and cancer. *Cell Death And Differentiation,* 17(2) pp. 209 - 214.

Luo, C. et al., 2012. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PLoS ONE,* 7(2) pp. e30087.

Luo, E.-C. et al., 2014. MicroRNA-769-3p down-regulates NDRG1 and enhances apoptosis in MCF-7 cells during reoxygenation. *Scientific Reports,* 4 pp. e5908.

Luo, G. et al., 2013. Highly lymphatic metastatic pancreatic cancer cells possess stem cell-like properties. *International Journal of Oncology*, 42(3) pp. 979 - 984.

Lv, D. et al., 2017. Three-dimensional cell culture: A powerful tool in tumor research and drug discovery. *Oncology letters*, 14(6) pp. 6999 - 7010.

Lv, J. et al., 2014. miRNA expression patterns in chemoresistant breast cancer tissues. *Biomedicine and Pharmacotherapy*, 68(8) pp. 935 - 942.

Lytle, N.K., Barber, A.G. and Reya, T., 2018. Stem cell fate in cancer growth, progression and therapy resistance. *Nature Reviews Cancer*, 18(11) pp. 669 - 680.

Ma, L., 2010. Role of miR-10b in breast cancer metastasis. *Breast Cancer Research*, 12(5) pp. 210 - 215.

Ma, R., Gong, J. and Jiang, X., 2017. Novel applications of next-generation sequencing in breast cancer research. *Genes and Diseases*, 4(3) pp. 149 - 153.

Madanecki, P. et al., 2013. Regulation of angiogenesis by hypoxia: the role of microRNA. *Cellular and Molecular Biology Letters*, 18(1) pp. 47 - 57.

Malik, K. and Brown, K.W., 2000. Epigenetic gene deregulation in cancer. *British Journal of Cancer*, 83(12) pp. 1583 - 1588.

Mani, S.A. et al., 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, 133(4) pp. 704 - 715.

Marcial-Quino, J. et al., 2016. Stem-loop RT-qPCR as an efficient tool for the detection and quantification of small RNAs in Giardia lamblia. *Genes*, 7(12) pp. 131 -146.

Mariel E. Toledo-Guzmán, Miguel Ibañez Hernández, Ángel A. Gómez-Gallegos and Ortiz-Sánchez, E., 2019. ALDH as a stem cell marker in solid tumors. *Current Stem Cell Research and Therapy*, 14 pp. e375.

Martello, G. et al., 2010. A microRNA targeting Dicer for metastasis control. *Cell*, 141(7) pp. 1195 - 1207.

Martin, T.A., Harrison, G., Mansel, R.E. and Jiang, W.G., 2003. The role of the CD44/ezrin complex in cancer metastasis. *Critical reviews in oncology/hematology*, 46(2) pp. 165 - 186.

Matter, M.S., Decaens, T., Andersen, J.B. and Thorgeirsson, S.S., 2014. Targeting the mTOR pathway in hepatocellular carcinoma: current state and future trends. *Journal of Hepatology*, 60(4) pp. 855 - 865.

Max, R. and Hannah, R. 2018. *Cancer* [*online*]. Available at: https://ourworldindata.org/cancer [Accessed 18 October 2019.

Maximiano, S., Magalhães, P., Guerreiro, M.P. and Morgado, M., 2016. Trastuzumab in the treatment of breast cancer. *BioDrugs,* 30(2) pp. 75 - 86.

Mazar, J. et al., 2016. MicroRNA 211 functions as a metabolic switch in human melanoma cells. *Molecular and cellular biology,* 36(7) pp. 1090 - 1108.

McCann, K.E. and Hurvitz, S.A., 2018. Advances in the use of PARP inhibitor therapy for breast cancer. *Drugs in Context,* 7 pp. 212540.

McVeigh, T.P. et al., 2014. The impact of Oncotype DX testing on breast cancer management and chemotherapy prescribing patterns in a tertiary referral centre. *European Journal of Cancer,* 50(16) pp. 2763 - 2770.

Mehta, P. et al., 2018. Self-renewal and CSCs *in vitro* enrichment: growth as floating spheres. *Methods in Molecular Biology,* 1692 pp. 61 - 75.

Meng, X. et al., 2015. Myeloid-specific TGF-β signaling in bone promotes basic-FGF and breast cancer bone metastasis. *Oncogene,* 35(18) pp. 2370 - 2378.

Merino, V.F. et al., 2016. Combined treatment with epigenetic, differentiating, and chemotherapeutic agents cooperatively targets tumor-Initiating cells in triple-negative breast cancer. *Cancer Research,* 76(7) pp. 2013 - 2024.

Meza-Sosa, K.F. et al., 2014. MiR-7 promotes epithelial cell transformation by targeting the tumor suppressor KLF4. *PLoS ONE,* 9(9) pp. e103987.

Mills, J.N., Rutkovsky, A.C. and Giordano, A., 2018. Mechanisms of resistance in estrogen receptor positive breast cancer: overcoming resistance to tamoxifen/aromatase inhibitors. *Current Opinion in Pharmacology,* 41 pp. 59 - 65.

Mishra, K., 2018. Intracellular reactive oxygen species determine cancer stem cell radiosensitivity related to predictive biomarker for radiotherapy. *Journal of Radiation and Cancer Research,* 9(4) pp. 147 - 154.

Misso, G. et al., 2014. Mir-34: a new weapon against cancer? *Molecular therapy. Nucleic acids,* 3(9) pp. e194.

Mo, J.-S. et al., 2015. MicroRNA 196B regulates FAS-mediated apoptosis in colorectal cancer cells. *Oncotarget,* 6(5) pp. 2843 - 2855.

Mollainezhad, H. et al., 2016. Expression of microRNA-370 in human breast cancer compare with normal samples. *Advanced Biomedical Research,* 5 pp. 129.

Moore, N. and Lyle, S., 2011. Quiescent, slow-cycling stem cell populations in cancer: a review of the evidence and discussion of significance. *Journal of Oncology,* 2011 pp. e396076.

Morel, A.-P. et al., 2008. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE,* 3(8) pp. e2888.

Mosmann, T., 1983. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *Journal of Immunological Methods,* 65(1) pp. 55 - 63.

Mu, J. et al., 2017. The repressive effect of miR-148a on Wnt/β-catenin signaling involved in Glabridin-induced anti-angiogenesis in human breast cancer cells. *BioMed Central Cancer,* 17(1) pp. 307.

Murali, V.S. et al., 2019. An image-based assay to quantify changes in proliferation and viability upon drug treatment in 3D microenvironments. *BioMed Central Cancer,* 19(1) pp. e502.

Murallitharan, M., 2018, Reasons why Malaysia has high cancer mortality rate [Online]. Available at: https://www.malaymail.com/news/malaysia/2018/12/13/reasons-why-malaysia-has-high-cancer-mortality-rate/1702830 [Accessed: 2 February 2019]

Nagelkerke, A., Bussink, J., Sweep, F.C.G.J. and Span, P.N., 2013. Generation of multicellular tumor spheroids of breast cancer cells: How to go three-dimensional. *Analytical Biochemistry,* 437(1) pp. 17 - 19.

Naito, H. et al., 2016. Endothelial side population cells contribute to tumor angiogenesis and antiangiogenic drug resistance. *Cancer Research,* 76(11) pp. 3200.

Najafi, M., Farhood, B. and Mortezaee, K., 2019. Cancer stem cells (CSCs) in cancer progression and therapy. *Journal of Cellular Physiology,* 234(6) pp. 8381 - 8395.

Nakai, K., Hung, M.-C. and Yamaguchi, H., 2016. A perspective on anti-EGFR therapies targeting triple-negative breast cancer. *American Journal of Cancer Research,* 6(8) pp. 1609 - 1623.

Nam, E. et al., 2012. MicroRNA profiling of a CD133+ spheroid-forming subpopulation of the OVCAR3 human ovarian cancer cell line. *BMC Medical Genomics,* 5(1) pp. e18.

Nam, R.K. et al., 2015. Identification and validation of a five microRNA signature predictive of prostate cancer recurrence and metastasis: a cohort study. *Journal of Cancer,* 6(11) pp. 1160 - 1171.

Narod, S.A. and Sopik, V., 2018. Is invasion a necessary step for metastases in breast cancer? *Breast Cancer Research and Treatment,* 169(1) pp. 9 - 23.

Nassar, F.J., Nasr, R. and Talhouk, R., 2017. MicroRNAs as biomarkers for early breast cancer diagnosis, prognosis and therapy prediction. *Pharmacology and Therapeutics,* 172 pp. 34 - 49.

Navarro, F. and Lieberman, J., 2015. MiR-34 and p53: new insights into a complex functional relationship. *PLoS ONE,* 10(7) pp. e0132767.

Nicolini, A., Ferrari, P. and Duffy, M.J., 2018. Prognostic and predictive biomarkers in breast cancer: Past, present and future. *Seminars in Cancer Biology,* 52 pp. 56 - 73.

Nunes, A.S. et al., 2018. 3D tumor spheroids as in vitro models to mimic *in vivo* human solid tumors resistance to therapeutic drugs. *Biotechnology and Bioengineering,* 116(1) pp. 206 - 226.

O'Brien, C.A., Kreso, A. and Dick, J.E., 2009. Cancer stem cells in solid tumors: an overview. *Seminars in Radiation Oncology,* 19(2) pp. 71 - 77.

O'Brien, J., Hayder, H., Zayed, Y. and Peng, C., 2018. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology,* 9 pp. 402.

Odoux, C. et al., 2008. A stochastic model for cancer stem cell origin in metastatic colon cancer. *Cancer Research,* 68(17) pp. 6932 - 6941.

Okamoto, K., et al., 2012. miR-493 induction during carcinogenesis blocks metastatic settlement of colon cancer cells in liver. *The EMBO journal*, 31(7), pp. 1752 – 1763.

Okuda, H. et al., 2013. miR-7 suppresses brain metastasis of breast cancer stem-like cells by modulating KLF4. *Cancer Research,* 73(4) pp. 1434 -1444.

Oztemur Islakoglu, Y., Noyan, S., Aydos, A. and Gur Dedeoglu, B., 2018. Meta-microRNA biomarker signatures to classify breast cancer subtypes. *OMICS: A Journal of Integrative Biology,* 22(11) pp. 709 - 716.

Pan, Y. et al., 2018. Therapeutic approaches targeting cancer stem cells. *Journal of Cancer Research and Therapeutics,* 14(7) pp. 1469 - 1475.

Pandit, H. et al., 2018. Enrichment of cancer stem cells via beta-catenin contributing to the tumorigenesis of hepatocellular carcinoma-catenin contributing to the tumorigenesis of hepatocellular carcinoma. *BioMed Cancer,* 18(1) pp. 783.

Panganiban, R.P. et al., 2019. Isolation and profiling of plasma microRNAs: Biomarkers for asthma and allergic rhinitis. *Methods,* 152 pp. 48 - 54.

Paoli, P., Giannoni, E. and Chiarugi, P., 2013. Anoikis molecular pathways and its role in cancer progression. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1833(12) pp. 3481 - 3498.

Paquette, M., El-Houjeiri, L. and Pause, A., 2018. mTOR pathways in cancer and autophagy. *Cancers,* 10(1) pp. 18.

Park, S.-M., Gaur, A.B., Lengyel, E. and Peter, M.E., 2008. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-

cadherin repressors ZEB1 and ZEB2. *Genes and Development,* 22(7) pp. 894 - 907.

Patani, N. and Martin, L.A., 2014. Understanding response and resistance to oestrogen deprivation in ER-positive breast cancer. *Molecular and Cellular Endocrinology,* 382(1) pp. 683 - 694.

Patel, N. et al., 2018. Regulating BMI1 expression via miRNAs promote Mesenchymal to Epithelial Transition (MET) and sensitizes breast cancer cell to chemotherapeutic drug. *PLoS ONE,* 13(2) pp. e0190245.

Pattabiraman, D.R. and Weinberg, R.A., 2014. Tackling the cancer stem cells-what challenges do they pose? *Nature Reviews. Drug Discovery,* 13(7) pp. 497 - 512.

Paul, P. et al., 2018. Interplay between miRNAs and human diseases. *Journal of Cellular Physiology,* 233(3) pp. 2007 - 2018.

Pearce, D.J. et al., 2005. Characterization of cells with a high aldehyde dehydrogenase activity from cord blood and acute myeloid leukemia samples. *Stem Cells,* 23(6) pp. 752 - 760.

Pece, S. et al., 2007. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell,* 140(1) pp. 62 - 73.

Peng, Y. and Croce, C.M., 2016. The role of microRNAs in human cancer. *Signal Transduction And Targeted Therapy,* 1 pp. 15004.

Peng, Y. et al., 2013. Insulin growth factor signaling is regulated by microRNA-486, an underexpressed microRNA in lung cancer. *Proceedings of the National Academy of Sciences,* 110(37) pp. 15043 - 15048.

Pereira, P.M. et al., 2010. MicroRNA expression variability in human cervical tissues. *PLoS ONE,* 5(7) pp. e11780.

Phi, L.T.H. et al., 2018. Cancer stem cells (CSCs) in drug resistance and their therapeutic implications in cancer treatment. *Stem Cells International,* 2018 pp. e16.

Phillips, T.M., McBride, W.H. and Pajonk, F., 2006. The response of CD24-/low/CD44+ breast cancer-initiating cells to radiation. *Journal of the National Cancer Institute,* 98(24) pp. 1777 - 1785.

Pickl, M. and Ries, C.H., 2008. Comparison of 3D and 2D tumor models reveals enhanced HER2 activation in 3D associated with an increased response to trastuzumab. *Oncogene,* 28(3) pp. 461 - 468.

Piovan, C. et al., 2012. Oncosuppressive role of p53-induced miR-205 in triple negative breast cancer. *Molecular Oncology,* 6(4) pp. 458 - 472.

Pópulo, H., Lopes, J.M. and Soares, P., 2012. The mTOR signalling pathway in human cancer. *International Journal of Molecular Sciences,* 13(2) pp. 1886 - 1918.

Porkka, K.P. et al., 2007. MicroRNA expression profiling in prostate cancer. *Cancer Research,* 67(13) pp. 6130 - 6135.

Prieto-Vila, M. et al., 2017. Drug resistance driven by cancer stem cells and their niche. *International Journal of Molecular Sciences,* 18(12) pp. 2574.

Pritchard, C.C., Cheng, H.H. and Tewari, M., 2012. MicroRNA profiling: approaches and considerations. *Nature Reviews. Genetics,* 13(5) pp. 358 - 369.

Pruszak, J. et al., 2009. CD15, CD24, and CD29 define a surface biomarker code for neural lineage differentiation of stem cells. *Stem Cells,* 27(12) pp. 2928 - 2940.

Punzi, S. et al., 2019. Development of personalized therapeutic strategies by targeting actionable vulnerabilities in metastatic and chemotherapy-resistant breast cancer PDXs. *Cells,* 8(6) pp. e605.

Qin, Q., Furong, W. and Baosheng, L., 2014. Multiple functions of hypoxia-regulated miR-210 in cancer. *Journal of Experimental and Clinical Cancer Research,* 33(1) pp. e50.

Quan, Y., Huang, X. and Quan, X., 2018. Expression of miRNA-206 and miRNA-145 in breast cancer and correlation with prognosis. *Oncology letters,* 16(5) pp. 6638 - 6642.

Radojicic, J. et al., 2011. MicroRNA expression analysis in triple-negative (ER, PR and Her2/neu) breast cancer. *Cell Cycle,* 10(3) pp. 507 - 517.

Rahimi, M. et al., 2019. An integrated analysis to predict micro-RNAs targeting both stemness and metastasis in breast cancer stem cells. *Journal of Cellular and Molecular Medicine,* 23(4) pp. e2442

Raisch J, Darfeuille-Michaud A and HTT., N., 2013. Role of microRNAs in the immune system, inflammation and cancer. *World journal of gastroenterology,* 19(20) pp. 2985 - 2996.

Ramgolam, K. et al., 2011. Melanoma spheroids grown under neural crest cell conditions are highly plastic migratory/invasive tumor cells endowed with immunomodulator function. *PLoS ONE,* 6(4) pp. e18784.

Rampersad, S.N., 2012. Multiple applications of Alamar Blue as an indicator of metabolic function and cellular health in cell viability bioassays. *Sensors (Basel, Switzerland),* 12(9) pp. 12347 - 12360.

Ray, A., Morford, R.K. and Provenzano, P.P., 2018. Cancer stem cell migration in three-dimensional aligned collagen matrices. *Current Protocols in Stem Cell Biology,* 46(1) pp. e57.

Ray, A., Vasudevan, S. and Sengupta, S., 2015. 6-Shogaol inhibits breast cancer cells and stem cell-like spheroids by modulation of Notch signaling pathway and induction of autophagic cell death. *PLoS ONE,* 10(9) pp. e0137614.

Redza-Dutordoir, M. and Averill-Bates, D.A., 2016. Activation of apoptosis signalling pathways by reactive oxygen species. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1863(12) pp. 2977 - 2992.

Ren, D. et al., 2017. Maintenance of cancer stemness by miR-196b-5p contributes to chemoresistance of colorectal cancer cells via activating STAT3 signaling pathway. *Oncotarget,* 8(30) pp. 49807 - 49823.

Ren, H., Qi, Y., Yin, X. and Gao, J., 2017. miR-136 targets MIEN1 and involves the metastasis of colon cancer by suppressing epithelial-to-mesenchymal transition. *OncoTargets and therapy,* 11 pp. 67 - 74.

Reynolds, B.A. and Weiss, S., 1996. Clonal and population analyses demonstrate that an EGF-responsive mammalian embryonic CNS precursor is a stem cell. *Developmental Biology,* 175(1) pp. 1 - 13.

Ribatti, D. et al., 2016. Angiogenesis and antiangiogenesis in triple-negative breast cancer. *Translational Oncology,* 9(5) pp. 453 - 457.

Ricardo, S. et al., 2011. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *Journal of Clinical Pathology,* 64(11) pp. 937 - 946.

Rich, J.N., 2016. Cancer stem cells: understanding tumor hierarchy and heterogeneity. *Medicine,* 95(1 Suppl 1) pp. 2 - 7.

Riffle, S., Pandey, R.N., Albert, M. and Hegde, R.S., 2017. Linking hypoxia, DNA damage and proliferation in multicellular tumor spheroids. *BioMed Central Cancer,* 17(1) pp. 338.

Roato, I. and Ferracini, R., 2018. Cancer stem cells, bone and tumor microenvironment: key players in bone metastases. *Cancers,* 10(2) pp. e56.

Robasky, K., Lewis, N.E. and Church, G.M., 2014. The role of replicates for error mitigation in next-generation sequencing. *Nature reviews. Genetics,* 15(1) pp. 56 - 62.

Roesch, A. et al., 2010. A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell,* 141(4) pp. 583 - 594.

Romero-Laorden, N. and Castro, E., 2017. Inherited mutations in DNA repair genes and cancer risk. *Current Problems in Cancer,* 41(4) pp. 251 - 264.

Roscigno, G. et al., 2017. MiR-24 induces chemotherapy resistance and hypoxic advantage in breast cancer. *Oncotarget,* 8(12) pp. 19507 - 19521.

Rose, C. et al., 2003. An open randomised trial of second-line endocrine therapy in advanced breast cancer: comparison of the aromatase inhibitors letrozole and anastrozole. *European Journal of Cancer,* 39(16) pp. 2318 - 2327.

Rothé , F. et al., 2011. Global microRNA expression profiling identifies miR-210 associated with tumor proliferation, invasion and poor clinical outcome in breast cancer. *PLoS ONE,* 6(6) pp. e20980.

Rugo, H.S. et al., 2016. Endocrine therapy for hormone receptor-positive metastatic breast cancer: American Society of Clinical Oncology Guideline. *Journal of Clinical Oncology,* 34(25) pp. 3069 - 3103.

Russo, J. and Russo, I.H., 1996. Experimentally induced mammary tumors in rats. *Breast Cancer Research and Treatment,* 39 pp. 7 - 20.

Sabry, D., Mostafa, A. and Hassouna, A., 2018. Breast carcinoma is a multifactorial disease involving FOXN3, SINA3 and NEAT through repression of GATA3 and TJP. *Journal of Thoracic Disease,* 10(3) pp. 1167 - 1171.

Saggar, J.K., Yu, M., Tan, Q. and Tannock, I.F., 2013. The tumor microenvironment and strategies to improve drug distribution. *Frontiers in Oncology,* 3 pp. 154 - 161.

Sánchez-Cid, L. et al., 2017. MicroRNA-200, associated with metastatic breast cancer, promotes traits of mammary luminal progenitor cells. *Oncotarget,* 8(48) pp. 83384 - 83406.

Sancho, P., Barneda, D. and Heeschen, C., 2016. Hallmarks of cancer stem cell metabolism. *British Journal of Cancer,* 114(12) pp. 1305 - 1312.

Sant, S. and Johnston, P.A., 2017. The production of 3D tumor spheroids for cancer drug discovery. *Drug discovery today. Technologies,* 23 pp. 27 - 36.

Sarkar, D. et al., 2008. Quality assessment and data analysis for microRNA expression arrays. *Nucleic Acids Research,* 37(2) pp. e17.

Scheel, C. and Weinberg, R.A., 2012. Cancer stem cells and epithelial to mesenchymal transition: concepts and molecular links. *Seminars in Cancer Biology,* 22(5-6) pp. 396-403.

Schober, M. and Fuchs, E., 2011. Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF-β and integrin/focal adhesion kinase (FAK) signaling. *Proceedings of the National Academy of Sciences of the United States of America,* 108(26) pp. 10544 - 10549.

Schroeder, A. et al., 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology,* 7 pp. 3 - 17.

Schulz, A., Meyer, F., Dubrovska, A. and Borgmann, K.C., 2019. Cancer stem cells and radioresistance: DNA repair and beyond. *Cancers,* 11(6) pp. e862

Schwarzenbacher, D., Balic, M. and Pichler, M., 2013. The role of microRNAs in breast cancer stem cells. *International Journal of Molecular Sciences,* 14(7) pp. 14712 - 14723.

Scully, T. et al., 2016. Insulin-like growth factor binding protein-3 links obesity and breast cancer progression. *Oncotarget,* 7(34) pp. 55491 - 55505.

Seely, J.M. and Alhassan, T., 2018. Screening for breast cancer in 2018 - what should we be doing today? *Current Oncology,* 25(Suppl 1) pp. 115 - 124.

Setijono, S.R. et al., 2018. miR-218 and miR-129 regulate breast cancer progression by targeting Lamins. *Biochemical and Biophysical Research Communications,* 496(3) pp. 826 - 833.

Shabani Azim, F., Houri, H., Ghalavand, Z. and Nikmanesh, B., 2018. Next generation sequencing in clinical oncology: applications, challenges and promises: a Review Article. *Iranian journal of public health,* 47(10) pp. 1453 - 1457.

Shah, M. and Calin, G., 2011. MicroRNAs miR-221 and miR-222: a new level of regulation in aggressive breast cancer. *Genome Medicine,* 3(8) pp. 56 - 60.

Shaheen, S., Ahmed, M., Lorenzi, F. and Nateri, A.S., 2016. Spheroid-formation (colonosphere) assay for *in vitro* assessment and expansion of stem cells in colon cancer. *Stem Cell Reviews and Reports,* 12(4) pp. 492 - 499.

Shakiba, N. et al., 2015. CD24 tracks divergent pluripotent states in mouse and human cells. *Nature Communications,* 6 pp. e7329.

Shan, X. et al., 2017. miR 1296-5p inhibits the migration and invasion of gastric cancer cells by repressing ERBB2 expression. *PLoS ONE,* 12(1) pp. e0170298.

Shao, C. et al., 2014. Essential role of aldehyde dehydrogenase 1A3 for the maintenance of non-small cell lung cancer stem cells is associated with the STAT3 pathway. *Clinical Cancer Research,* 20(15) pp. 4154 - 4166.

Sharp, J.W. et al., 2019. Modifiable patient-related barriers and their association with breast cancer detection practices among Ugandan women without a diagnosis of breast cancer. *PLoS ONE,* 14(6) pp. e0217938.

Shen, S. et al., 2017. miR-204 regulates the biological behavior of breast cancer MCF-7 cells by directly targeting FOXA1. *Oncology Reports,* 38(1) pp. 368 - 376.

Shen, S. et al., 2013. Upregulation of miR-136 in human non-small cell lung cancer cells promotes Erk1/2 activation by targeting PPP2R2A. *Tumor Biology,* 35(1) pp. 631 - 640.

Sheng, W. et al., 2016. miR-148a inhibits self-renewal of thyroid cancer stem cells via repressing INO80 expression. *Oncology Reports,* 36(6) pp. 3387 - 3396.

Sheridan, C. et al., 2006. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Research,* 8(5) pp. R59.

Shi, L. et al., 2017. miR-127 promotes EMT and stem-like traits in lung cancer through a feed-forward regulatory loop. *Oncogene,* 36 pp. 1631.

Shi, W.-l. et al., 2016. Integrated miRNA and mRNA expression profiling in fetal hippocampus with Down syndrome. *Journal of Biomedical Science,* 23(1) pp. 48 - 92.

Shi, W. et al., 2011. MicroRNA-301 mediates proliferation and invasion in human breast cancer. *Cancer Research,* 71(8) pp. 2926 - 2937.

Shi, X., Zhang, Y., Zheng, J. and Pan, J., 2012. Reactive oxygen species in cancer stem cells. *Antioxidants and Redox Signaling,* 16(11) pp. 1215 - 1228.

Shi, Y. et al., 2014. A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). *Journal of Thoracic Cncology,* 9(2) pp. 154 - 162.

Shi Y, Ye P and Long X., 2017. Differential expression profiles of the transcriptome in breast cancer cell lines revealed by next generation sequencing. *Cellular Physiology and Biochemistry,* 44 pp. 804 – 816.

Shibue, T. and Weinberg, R.A., 2017. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nature Reviews. Clinical Oncology,* 14(10) pp. 611 - 629.

Shibuya, K. et al., 2015. Targeting the facilitative glucose transporter GLUT1 inhibits the self-renewal and tumor-initiating capacity of cancer stem cells. *Oncotarget,* 6(2) pp. 651 - 661.

Shimono, Y., Mukohyama, J., Nakamura, S.-I. and Minami, H., 2016. MicroRNA regulation of human breast cancer stem cells. *Journal of Clinical Medicine,* 5(1) pp. 1 - 24.

Shimono, Y. et al., 2009. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell,* 138(3) pp. 592 - 603.

Shmelkov, S.V. et al., 2008. CD133 expression is not restricted to stem cells, and both CD133+ and CD133- metastatic colon cancer cells initiate tumors. *The Journal of Clinical Investigation,* 118(6) pp. 2111 - 2120.

Shore, S. et al., 2016. Small RNA library preparation method for next-generation sequencing using chemical modifications to prevent adapter dimer formation. *PLoS ONE,* 11(11) pp. e0167009.

Sidney, L.E. et al., 2014. Concise review: evidence for CD34 as a common marker for diverse progenitors. *Stem Cells,* 32(6) pp. 1380 - 1389.

Siegel, R.L., Miller, K.D. and Jemal, A., 2018. Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians,* 68(1) pp. 7 - 30.

Siegel, R.L., Miller, K.D. and Jemal, A., 2019. Cancer statistics, 2019. *CA: A Cancer Journal for Clinicians,* 69(1) pp. 7 - 34.

Sledge, G.W. et al., 2003. Phase III Trial of doxorubicin, paclitaxel, and the combination of doxorubicin and paclitaxel as front-line chemotherapy for metastatic breast cancer: an intergroup trial (E1193). *Journal of Clinical Oncology,* 21(4) pp. 588 - 592.

Sokolenko, A.P. and Imyanitov, E.N., 2018. Molecular diagnostics in clinical oncology. *Frontiers in molecular biosciences,* 5 pp. 76.

Song, Y. et al., 2017. A unified model of the hierarchical and stochastic theories of gastric cancer. *British Journal of Cancer,* 116(8) pp. 973 - 989.

Sorlie, T. et al., 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences,* 98(19) pp. 10869 - 10874.

Stefanie Sassen, Eric A. Miska and Caldas, C., 2008. MicroRNA—implications for cancer. *European Journal of Pathology,* 452(1) pp. 1 - 10.

Stinson, S. et al., 2011. TRPS1 targeting by miR-221/222 promotes the epithelial-to-mesenchymal transition in breast cancer. *Science Signal,* 4(177) pp. ra41.

Sun, C. et al., 2015. Hsa-miR-139-5p inhibits proliferation and causes apoptosis associated with down-regulation of c-Met. *Oncotarget,* 6(37) pp. 39756 - 39792.

Sun, C. et al., 2018. Diagnostic and prognostic value of microRNAs in cholangiocarcinoma: a systematic review and meta-analysis. *Cancer Management and Research,* 10 pp. 2125 - 2139.

Sun, Y.-S. et al., 2017. Risk factors and preventions of breast cancer. *International Journal of Biological Sciences,* 13(11) pp. 1387 - 1397.

Suzuki, H., Maruyama, R., Yamamoto, E. and Kai, M., 2012. DNA methylation and microRNA dysregulation in cancer. *Molecular Oncology,* 6(6) pp. 567 - 578.

Svoboda, P., 2015. A toolbox for miRNA analysis. *Federation of European Biochemical Societies Letters,* 589(14) pp. 1694 -1701.

Swellam, M., Zahran, R.F.K., Ghonem, S.A. and Abdel-Malak, C., 2019. Serum miRNA-27a as potential diagnostic nucleic marker for breast cancer. *Archives of Physiology and Biochemistry,* pp. 1 - 7.

Takahashi, R.-u., Miyazaki, H. and Ochiya, T., 2014. The role of microRNAs in the regulation of cancer stem cells. *Frontiers in Genetics,* 4 pp. 1-11.

Takebe, N. et al., 2015. Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: clinical update. *Nature Reviews. Clinical oncology.,* 12(8) pp. 445 - 464.

Takeishi, S. and Nakayama, K.I., 2016. To wake up cancer stem cells, or to let them sleep, that is the question. *Cancer science,* 107(7) pp. 875 - 881.

Tam, S., de Borja, R., Tsao, M.-S. and McPherson, J.D., 2014. Robust global microRNA expression profiling using next-generation sequencing technologies. *Laboratory Investigation,* 94(3) pp. 350 - 358.

Tamura, M. et al., 2017. Morphology-based optical separation of subpopulations from a heterogeneous murine breast cancer cell line. *PLoS ONE,* 12(6) pp. e0179372.

Tan, W. et al., 2018. MicroRNAs and cancer: Key paradigms in molecular therapy. *Oncology Letters,* 15(3) pp. 2735 - 2742.

Tan, Y. et al., 2009. Transcriptional inhibiton of Hoxd4 expression by miRNA-10a in human breast cancer cells. *BMC Molecular Biology,* 10(1) pp. e12.

Tancioni, I. et al., 2015. FAK activity protects nucleostemin in facilitating breast cancer spheroid and tumor growth. *Breast Cancer Research,* 17(1) pp. 47 - 63.

Taniguchi, H., Suzuki, Y. and Natori, Y., 2019. The evolving landscape of cancer stem cells and ways to overcome cancer heterogeneity. *Cancers,* 11(4) pp. 532 - 548.

Taylor, M.A. et al., 2012. TGF-B upregulates miR-181a expression to promote breast cancer metastasis. *The Journal of Clinical Investigation,* 123(1) pp. 150 - 163.

Tazawa, H., Tsuchiya, N., Izumiya, M. and Nakagama, H., 2007. Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. *Proceedings of the National Academy of Sciences of the United States of America,* 104(39) pp. 15472 - 15477.

Tevis, K.M., Colson, Y.L. and Grinstaff, M.W., 2017. Embedded spheroids as models of the cancer microenvironment. *Advanced Biosystems,* 1(10) pp. 1700083.

Thomas, M., Kelly, E.D., Abraham, J. and Kruse, M., 2019. Invasive lobular breast cancer: A review of pathogenesis, diagnosis, management, and future directions of early stage disease. *Seminars in Oncology,* 46(2) pp. 121-132.

Tirino, V. et al., 2013. Cancer stem cells in solid tumors: an overview and new approaches for their isolation and characterization. *The FASEB Journal,* 27(1) pp. 13 - 24.

To, K.K.W., 2013. MicroRNA: a prognostic biomarker and a possible druggable target for circumventing multidrug resistance in cancer chemotherapy. *Journal of Biomedical Science,* 20(1) pp. 99 - 118.

Tong, C.W.S., Wu, M., Cho, W.C.S. and To, K.K.W., 2018. Recent advances in the treatment of breast cancer. *Frontiers in Oncology,* 8 pp. 227.

Torre, L.A. et al., 2017. Global cancer in women: burden and trends. *Cancer Epidemiology, Biomarkers and Prevention,* 26(4) pp. 444.

Tredan, O., Galmarini, C.M., Patel, K. and Tannock, I.F., 2007. Drug resistance and the solid tumor microenvironment. *Journal of the National Cancer Institute,* 99(19) pp. 1441 - 1454.

Treiber, T., Treiber, N. and Meister, G., 2019. Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nature Reviews Molecular Cell Biology,* 20(1) pp. 5 - 20.

Tremont, A., Lu, J. and Cole, J.T., 2017. Endocrine therapy for early breast cancer: updated review. *The Ochsner journal,* 17(4) pp. 405 - 411.

Tubbs, A. and Nussenzweig, A., 2017. Endogenous DNA damage as a source of genomic instability in cancer. *Cell,* 168(4) pp. 644 - 656.

Turdo, A. et al., 2019. Meeting the challenge of targeting cancer stem cells. *Frontiers in cell and developmental biology,* 7(16) pp. e30834247.

Turner, N.D. and Lloyd, S.K., 2017. Association between red meat consumption and colon cancer: A systematic review of experimental results. *Experimental Biology and Medicine,* 242(8) pp. 813 - 839.

Vahidian, F. et al., 2018. MicroRNAs and breast cancer stem cells: Potential role in breast cancer therapy. *Journal of Cellular Physiology,* 234(4) pp. 3294 - 3306.

Valentina, I. et al., 2016. One hundred faces of cyclopamine. *Current Pharmaceutical Design,* 22(12) pp. 1658 - 1681.

van Niekerk, G., Davids, L.M., Hattingh, S.M. and Engelbrecht, A.-M., 2017. Cancer stem cells: A product of clonal evolution? *International Journal of Cancer,* 140(5) pp. 993 - 999.

van Schooneveld, E. et al., 2012. Expression profiling of cancerous and normal breast tissues identifies microRNAs that are differentially expressed in serum from patients with (metastatic) breast cancer and healthy volunteers. *Breast Cancer Research,* 14(1) pp. R34.

Vannini, I., Fanini, F. and Fabbri, M., 2018. Emerging roles of microRNAs in cancer. *Current Opinion in Genetics and Development,* 48 pp. 128 - 133.

Vassalli, G., 2019. Aldehyde dehydrogenases: not just markers, but functional regulators of stem cells. *Stem Cells International,* 2019 pp. e3904645.

Vermeersch, K.A. et al., 2015. OVCAR-3 spheroid-derived cells display distinct metabolic profiles. *PLoS ONE,* 10(2) pp. e0118262.

Vishnubalaji, R. et al., 2015. Genome-wide mRNA and miRNA expression profiling reveal multiple regulatory networks in colorectal cancer. *Cell Death and Disease,* 6 pp. e1614.

Vishnubalaji, R. et al., 2018. Molecular profiling of ALDH1(+) colorectal cancer stem cells reveals preferential activation of MAPK, FAK, and oxidative stress pro-survival signalling pathways. *Oncotarget,* 9(17) pp. 13551 - 13564.

Vlachos, I.S. et al., 2015. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Research,* 43(W1) pp. W460 - W466.

Vogelstein, B. et al., 2013. Cancer genome landscapes. *Science,* 339(6127) pp. 1546 - 1558.

Volonte, D. et al., 2018. Caveolin-1 promotes the tumor suppressor properties of oncogene-induced cellular senescence. *Journal of Biological Chemistry,* 293(5) pp. 1794 - 1809.

Voorhoeve, P.M., 2010. MicroRNAs: oncogenes, tumor suppressors or master regulators of cancer heterogeneity? *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer,* 1805(1) pp. 72 - 86.

Waks, A.G. and Winer, E.P., 2019. Breast cancer treatment: A review *Journal of the American Medical Association,* 321(3) pp. 288 - 300.

Walsh, M.F., Nathanson, K.L., Couch, F.J. and Offit, K., 2016. Genomic biomarkers for breast cancer risk. *Advances in Experimental Medicine and Biology,* 882 pp. 1 - 32.

Wang, B. et al., 2011. Systematic evaluation of three microRNA profiling platforms: microarray, beads array, and quantitative real-time PCR array. *PLoS ONE,* 6(2) pp. e17167.

Wang, B. and Xi, Y., 2013. Challenges for microRNA microarray data analysis. *Microarrays,* 2(2) pp. 34 - 50.

Wang, F. et al., 2016. Hedgehog signaling regulates epithelial-mesenchymal transition in pancreatic cancer stem-like cells. *Journal of Cancer,* 7(4) pp. 408 - 417.

Wang, H. et al., 2017. Investigating microRNA and transcription factor co-regulatory networks in colorectal cancer. *Biomed Central Bioinformatics,* 18(1) pp. 388

Wang, J. et al., 2016. Transition to resistance: An unexpected role of the EMT in cancer chemoresistance. *Genes and Diseases,* 3(1) pp. 3 - 6.

Wang, K., Chen, M. and Wu, W., 2017. Analysis of microRNA (miRNA) expression profiles reveals 11 key biomarkers associated with non-small cell lung cancer. *World journal of surgical oncology,* 15(1) pp. 175.

Wang, L. et al., 2014. Enrichment and characterization of cancer stem-like cells from a cervical cancer cell line. *Molecular Medicine Reports,* 9(6) pp. 2117 - 2123.

Wang, R. et al., 2014. Comparison of mammosphere formation from breast cancer cell lines and primary breast tumors. *Journal of Thoracic Disease,* 6(6) pp. 829 - 837.

Wang, T., Zheng, L., Wang, Q. and Hu, Y.-W., 2018. Emerging roles and mechanisms of FOXC2 in cancer. *Clinica Chimica Acta,* 479 pp. 84 - 93.

Wang, X., Guo, H., Yao, B. and Helms, J., 2017. miR-15b inhibits cancer-initiating cell phenotypes and chemoresistance of cisplatin by targeting TRIM14 in oral tongue squamous cell cancer. *Oncology Reports,* 37(5) pp. 2720 - 2726.

Wang, Y. et al., 2011. Transforming growth factor- β regulates the sphere-initiating stem cell-like feature in breast cancer through miRNA-181 and ATM. *Oncogene,* 30(12) pp. 1470 - 1480.

Wang, Z. and Liu, C., 2016. MiR-153 regulates metastases of gastric cancer through Snail. *Tumor Biology,* 37(12) pp. 15509 - 15515.

Wang, Z. et al., 2019. The prognostic and clinical value of CD44 in colorectal cancer: A meta-analysis. *Frontiers in Oncology,* 9 pp. 309 - 309.

Wang, Z., et al., 2015. MicroRNA-378-5p suppresses cell proliferation and induces apoptosis in colorectal cancer cells by targeting BRAF. *Cancer Cell International*, 15 pp. e40.

Ward Rashidi, M.R. et al., 2019. Engineered 3D model of cancer stem cell enrichment and chemoresistance. *Neoplasia,* 21(8) pp. 822 - 836.

Wei, W. et al., 2012. Relationship of CD44+CD24-/low breast cancer stem cells and axillary lymph node metastasis. *Journal of Translational Medicine,* 10(Suppl 1) pp. S6.

Weiner, L.M., Surana, R. and Wang, S., 2010. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nature Reviews. Immunology,* 10(5) pp. 317 - 327.

Weinhold, N. et al., 2014. Genome-wide analysis of noncoding regulatory mutations in cancer. *Nature Genetics*, 46(11) pp. 1160 - 1165.

Weiswald, L.-B., Bellet, D. and Dangles-Marie, V., 2015. Spherical cancer models in tumor biology. *Neoplasia,* 17(1) pp. 1 - 15.

Welch, D.R. and Hurst, D.R., 2019. Defining the hallmarks of metastasis. *Cancer Research,* 79(12) pp. 3011 - 3027.

Wheler, J.J. et al., 2016. Presence of both alterations in FGFR/FGF and PI3K/AKT/mTOR confer improved outcomes for patients with metastatic breast cancer treated with PI3K/AKT/mTOR inhibitors. *Oncoscience,* 3(5-6) pp. 164 - 172.

White, J. and Dalton, S., 2005. Cell cycle control of embryonic stem cells. *Stem Cell Reviews,* 1(2) pp. 131 - 138.

Wilkinson, H.N. and Hardman, M.J., 2017. The role of estrogen in cutaneous ageing and repair. *Maturitas,* 103 pp. 60 - 64.

Wöckel, A. et al., 2018. The screening, diagnosis, treatment, and follow-ip of breast cancer. *Deutsches Arzteblatt international,* 115(18) pp. 316 - 323.

Wu, F., Yang, Z. and Li, G., 2009. Role of specific microRNAs for endothelial function and angiogenesis. *Biochemical and Biophysical Research Communications,* 386(4) pp. 549 - 553.

Wu, H. et al., 2014. Sphere formation assay is not an effective method for cancer stem cell derivation and characterization from the Caco-2 colorectal cell line. *Current Stem Cell Research and Therapy,* 9(2) pp. 82 - 88.

Wu, H., Zhu, S. and Mo, Y.-Y., 2009. Suppression of cell growth and invasion by miR-205 in breast cancer. *Cell research,* 19(4) pp. 439 - 448.

Wu, X., Li, L., Li, Y. and Liu, Z., 2016. MiR-153 promotes breast cancer cell apoptosis by targeting HECTD3. *American journal of cancer research,* 6(7) pp. 1563 - 1571.

X Wang et al., 2017. Differential expression profile analysis of miRNAs with HER-2 overexpression and intervention in breast cancer cells. *International Journal of Clinical and Experimental Pathology,* 10(5) pp. 5039 - 5062.

Xia, K. et al., 2015. miR-411 regulated ITCH expression and promoted cell proliferation in human hepatocellular carcinoma cells. *Biomedicine and Pharmacotherapy,* 70(0) pp. 158 - 163.

Xu, Q. et al., 2013. A regulatory circuit of miR-148a/152 and DNMT1 in modulating cell transformation and tumor angiogenesis through IGF-IR and IRS1. *Journal of molecular cell biology,* 5(1) pp. 3 - 13.

Xu, W. et al., 2016. HIF-2α, acting via miR-191, is involved in angiogenesis and metastasis of arsenite-transformed HBE cells. *Toxicology research,* 5(1) pp. 66 - 78.

Xu, Y.-F. et al., 2017. Plasma exosome miR-196a and miR-1246 are potential indicators of localized pancreatic cancer. *Oncotarget,* 8(44) pp. 77028 - 77040.

Xu, Y. et al., 2013. Changes in the expression of miR-381 and miR-495 are inversely associated with the expression of the MDR1 gene and development of multi-drug resistance. *PLoS ONE,* 8(11) pp. e82062.

Yamaguchi, T. et al., 2008. Activation of the intrinsic and extrinsic pathways in high pressure-induced apoptosis of murine erythroleukemia cells. *Cellular and Molecular Biology Letters,* 13(1) pp. 49 - 57.

Yamamoto, H. et al., 2012. CD49f is an efficient marker of monolayer- and spheroid colony-forming cells of the benign and malignant human prostate. *PLoS ONE,* 7(10) pp. e46979.

Yan H et al., 2015. Up-regulation of miR-204 enhances anoikis sensitivity in epithelial ovarian cancer cell line via brain-derived neurotrophic factor pathway *in vitro*. *International Journal of Gynecological Cancer,* 25(6) pp. 944 - 952.

Yan, L.X. et al., 2011. Knockdown of miR-21 in human breast cancer cell lines inhibits proliferation, in vitro migration and in vivo tumor growth. *Breast Cancer Research,* 13(1) pp. R2.

Yan, M. et al., 2016. miR-136 suppresses tumor invasion and metastasis by targeting RASAL2 in triple-negative breast cancer. *Oncology Reports,* 36 pp. 65 - 71.

Yan, W. et al., 2013. Increased invasion and tumorigenicity capacity of CD44+/CD24- breast cancer MCF7 cells in vitro and in nude mice. *Cancer Cell International,* 13(1) pp. 62 - 69.

Yan, Y., Zuo, X. and Wei, D., 2015. Concise review: emerging role of CD44 in cancer stem cells: a promising biomarker and therapeutic target. *Stem Cells Translational Medicine,* 4(9) pp. 1033 - 1043.

Yan, Z.-Y. et al., 2016. Integrated analysis and microRNA expression profiling identified seven miRNAs associated with progression of oral squamous cell carcinoma. *Journal of Cellular Physiology,* 232(8) pp. 2178 - 2185.

Yang, C.-H. et al., 2014. Identification of CD24 as a cancer stem cell marker in human nasopharyngeal carcinoma. *PLoS ONE,* 9(6) pp. e99412.

Yang, C., Tabatabaei, S.N., Ruan, X. and Hardy, P., 2017. The dual regulatorr role of miR-181a in breast cancer. *Cellular Physiology and Biochemistry,* 44(3) pp. 843 - 856.

Yang, Y. et al., 2015. Side population cells separated from A549 lung cancer cell line possess cancer stem cell-like properties and inhibition of autophagy potentiates the cytotoxic effect of cisplatin. *Oncology Reports,* 34 pp. 929 - 935.

Yang, Z. et al., 2018. Transcriptome profiling of Panc-1 spheroid cells with pancreatic cancer stem cells properties cultured by a novel 3D semi-solid system. *Cellular Physiology and Biochemistry,* 47(5) pp. 2109 - 2125.

Yaşar P, Ayaz G, User SD and Güpür G, M.M., 2016. Molecular mechanism of estrogen-estrogen receptor signaling. *Reproductive Medicine and Biology,* 16(1) pp. 4 - 20.

Yasuda, K. et al., 2013. Ovarian cancer stem cells are enriched inside population and aldehyde dehydrogenase bright overlapping population. *PLoS ONE,* 8(8) pp. e68187.

Yasuda, M.T., Sakakibara, H. and Shimoi, K., 2017. Estrogen- and stress-induced DNA damage in breast cancer and chemoprevention with dietary flavonoid. *Genes and Environment: the official journal of the Japanese Environmental Mutagen Society,* 39 pp. 10.

Yeap, S. K., 2010. Immunomodulatory effects of *Rhaphidophora Korthalsii* methanol extract on natural killer cell activation and cytolytic activity. PhD thesis, Universiti Putra Malaysia, Malaysia.

Yersal, O. and Barutca, S., 2014. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology,* 5(3) pp. 412 - 424.

Yim, E. and Sheetz, M., 2012. Force-dependent cell signaling in stem cell differentiation. *Stem Cell Research and Therapy,* 3(5) pp. e41.

Yin Y et al., 2014. miR-204-5p inhibits proliferation and invasion and enhances chemotherapeutic sensitivity of colorectal cancer cells by downregulating RAB22A. *Clinical Cancer Research,* 20(23) pp. 6187 - 6199.

Yoo, B. et al., 2017. Therapy targeted to the metastatic niche is effective in a model of stage IV breast cancer. *Scientific Reports,* 7 pp. 45060.

Yoo, J.-O. et al., 2016. miR-181b-3p promotes epithelial-mesenchymal transition in breast cancer cells through Snail stabilization by directly targeting YWHAG. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1863(7, Part A) pp. 1601 - 1611.

Yoshida, G.J. and Saya, H., 2016. Therapeutic strategies targeting cancer stem cells. *Cancer science,* 107(1) pp. 5 - 11.

You, Y. et al., 2018. Ovarian cancer stem cells promote tumour immune privilege and invasion via CCL5 and regulatory T cells. *Clinical and Experimental Immunology,* 191(1) pp. 60 - 73.

Yu, F. et al., 2012. MicroRNA 34c gene downregulation via DNA methylation promotes self-renewal and epithelial-mesenchymal transition in breast tumor-initiating cells. *The Journal of Biological Chemistry,* 287(1) pp. 465 - 473.

Yu, F. et al., 2007. let-7 regulates self-renewal and tumorigenicity of breast cancer cells. *Cell,* 131(6) pp. 1109 - 1123.

Yu, J. et al., 2016. MicroRNA-182 targets SMAD7 to potentiate TGFβ-induced epithelial-mesenchymal transition and metastasis of cancer cells. *Nature Communications,* 7 pp. 13884.

Yu, J. et al., 2011. MiR-148a inhibits angiogenesis by targeting ERBB3. *Journal of biomedical research,* 25(3) pp. 170 - 177.

Yu, Y. et al., 2018. miR-190 suppresses breast cancer metastasis by regulation of TGF-β-induced epithelial-mesenchymal transition. *Molecular Cancer,* 17(1) pp. e70

Yuan, F. and Wang, W., 2015. MicroRNA-802 suppresses breast cancer proliferation through downregulation of FoxM1. *Molecular Medicine Reports,* 12(3) pp. 4647 - 4651.

Yun, Z. and Lin, Q., 2014. Hypoxia and regulation of cancer cell stemness. *Advances in Experimental Medicine and Biology,* 772 pp. 41 - 53.

Zhang, C. et al., 2015. MiR-33a suppresses breast cancer cell proliferation and metastasis by targeting ADAM9 and ROS1. *Protein and Cell,* 6(12) pp. 881 - 889.

Zhang, H. et al., 2014. MiR-7, inhibited indirectly by LincRNA HOTAIR, directly inhibits SETDB1 and reverses the EMT of breast cancer stem cells by downregulating the STAT3 pathway. *STEM CELLS,* 32(11) pp. 2858 - 2868.

Zhang, J. et al., 2014. Genome-wide uncovering of STAT3-mediated miRNA expression profiles in colorectal cancer cell lines. *BioMed Research International,* 2014 pp. e187105.

Zhang, M.H. et al., 2014. Estrogen receptor-positive breast cancer molecular signatures and therapeutic potentials (Review). *Biomedical Report,* 2 pp. 41 - 52.

Zhang, W. and Zhang, J., 2015. miR-181c promotes proliferation via suppressing PTEN expression in inflammatory breast cancer. *International Journal of Oncology,* 46(5) pp. 2011 - 2020.

Zhang, X., Yang, D. and Wei, Y., 2018. Overexpressed CDR1as functions as an oncogene to promote the tumor progression via miR-7 in non-small-cell lung cancer. *OncoTargets and Therapy,* 11 pp. 3979 - 3987.

Zhao, C. et al., 2017. miR-15b-5p resensitizes colon cancer cells to 5-fluorouracil by promoting apoptosis via the NF-κB axis. *Scientific Reports,* 7(1) pp. 4194.

Zhao, H. et al., 2015. Expression of miR-136 is associated with the primary cisplatin resistance of human epithelial ovarian cancer. . *Oncology Reports,* 33 pp. 591 - 598.

Zhao, J., 2016. Cancer stem cells and chemoresistance: The smartest survives the raid. *Pharmacology and Therapeutics,* 160 pp. 145 - 158.

Zhao, L. et al., 2016. miR-493-5p attenuates the invasiveness and tumorigenicity in human breast cancer by targeting FUT4. *Oncology Reports,* 36 pp. 1007 - 1015.

Zhao, W., Li, Y. and Zhang, X., 2017. Stemness-related markers in cancer. *Cancer Translational Medicine,* 3(3) pp. 87 - 95.

Zhao, Y. et al., 2018. TGF-β transactivates EGFR and facilitates breast cancer migration and invasion through canonical Smad3 and ERK/Sp1 signaling pathways. *Molecular Oncology,* 12(3) pp. 305 - 321.

Zhao, Z. et al., 2017. miR-503-3p promotes epithelial-mesenchymal transition in breast cancer by directly targeting SMAD2 and E-cadherin. *Journal of Genetics and Genomics,* 44(2) pp. 75 - 84.

Zheng, J.-Z. et al., 2018. Elevated miR-301a expression indicates a poor prognosis for breast cancer patients. *Scientific Reports,* 8(1) pp. e2225.

Zhi, D., Zhao, X., Dong, M. and Yan, C., 2017. miR-493 inhibits proliferation and invasion in pancreatic cancer cells and inversely regulated hERG1 expression. *Oncology letters,* 14(6) pp. 7398 - 7404.

Zhou, J., Yi, Q. and Tang, L., 2019. The roles of nuclear focal adhesion kinase (FAK) on Cancer: a focused review. *Journal of Experimental & Clinical Cancer Research,* 38(1) pp. 1 - 12.

Zhou, W., et al., 2015. MiR-493 suppresses the proliferation and invasion of gastric cancer cells by targeting RhoC. *Iranian Journal of Basic Medical Sciences,* 18(10) pp. 1027 – 1033.

Zhou, X. et al., 2017. MicroRNA-145 inhibits tumorigenesis and invasion of cervical cancer stem cells. *International Journal of Oncology,* 50(3) pp. 853 - 862.

Zhu, Y.-T. et al., 2018. A modified method by differential adhesion and serum-free culture medium for enrichment of cancer stem cells. *Journal of Cancer Research and Therapeutics,* 14(9) pp. 421 - 426.

Zucca-Matthes, G., Urban, C. and Vallejo, A., 2016. Anatomy of the nipple and breast ducts. *Gland Surgery,* 5(1) pp. 32 - 36.

LIST OF PUBLICATIONS

First author

1. **Boo L**, Ho WY, Ali NM, et al. MiRNA Transcriptome Profiling of Spheroid-Enriched Cells with Cancer Stem Cell Properties in Human Breast MCF-7 Cell Line. *International Journal of Biological Sciences*. 2016;12(4):427-445. doi:10.7150/ijbs.12777. Tier 1, IF: 4.5

2. **Boo L**, Ho WY, Mohd Ali N, et al. Phenotypic and microRNA transcriptomic profiling of the MDA-MB-231 spheroid-enriched CSCs with comparison of MCF-7 microRNA profiling dataset. *PeerJ*. 2017;5: e3551. doi:10.7717/peerj.3551. Tier 2, IF: 2.2

3. **Boo L**, Yeap SK, Ali NM, Ho WY, Ky H, Satharasinghe DA, Liew WC, Tan SW, Wang M-L, Cheong SK, and Ong HK. 2019. Phenotypic and microRNA characterisation of the neglected CD24+ cell population in MCF-7 breast cancer 3D spheroid culture. *Journal of the Chinese Medical Association* Latest Articles. Tier 2, IF: 1.894

Co-authors

1. Mohd Ali N, **Boo L**, Yeap SK, Ky H, Satharasinghe DA, Liew WC, Ong HK, Cheong SK, and Kamarul T. 2016. Probable impact of age and hypoxia on proliferation and microRNA expression profile of bone marrow-derived human mesenchymal stem cells. PeerJ 4: e1536. 10.7717/peerj.1536. Tier 2, IF: 2.2

APPENDICES

APPENDIX A

Primers used for mycoplasma reactions

MYCO F1: 5’-CGC CTG AGT AGT ACG TWC GC-3’

MYCO F2: 5’-TGC CTG RGT AGT ACA TTC GC-3’

MYCO F3: 5’-CGC CTG AGT AGT ATG CTC GC-3’

MYCO F4: 5’-CGC CTG GGT AGT ACA TTC GC-3’

MYCO R1: 5’-GCG GTG TGT ACA ARA CCC GA-3’

MYCO R2: 5’-GCG GTG TGT ACA AAC CCC GA-3’

(R= MIXTURE OF G AND A; W= MIXTURE OF T AND A)

APPENDIX B

Spheroid-forming efficiency (SFE) of secondary spheroids of MCF-7 spheroid dissociated cells from first to third generation

Generation	Well	Number of tumourspheres	Sphere forming efficiency %	SD
First	1	57	28.5	
	2	45	22.5	
	3	44	22	
	Avg	52	24.3	3.62
	Well	Number of tumourspheres	Sphere forming efficiency %	
Second	1	56	28.1	
	2	57	28.4	
	3	57	28.4	
	Avg	57	28.3	0.17
	Well	Number of tumourspheres	Sphere forming efficiency %	
Third	1	45	22.5	
	2	42	21	
	3	39	19.5	
	Avg	42	21.0	1.50

Three independent experiments were carried out. (Equivalent to Figure 4.10A)

APPENDIX C

Spheroid-forming efficiency (SFE) of secondary spheroids of MDA-MB-231 spheroid dissociated cells from first to third generation

Generation	Well	Number of tumourspheres	Sphere forming efficiency %	SD
First	1	47	23.5	
	2	58	29	
	3	45	22.5	
	Avg	50	25.0	3.50
	Well	Number of tumourspheres	Sphere forming efficiency %	
Second	1	45	22.5	
	2	40	20	
	3	53	26.5	
	Avg	46	23.0	3.28
	Well	Number of tumourspheres	Sphere forming efficiency %	
Third	1	56	28	
	2	65	32.5	
	3	63	31.5	
	Avg	61	30.7	2.36

Three independent experiments were carried out. (Equivalent to Figure 4.10A)

APPENDIX D

Percentage of cells stained with CD44+/CD24- and ALDH+ in MCF-7 and MDA-MB-231 spheroid relative to their respective parental cells

% cells stained with CD44+/CD24-

	Replicates				
Cell type	1	2	3	Mean	SD
MCF7 Parental	0.30	0.43	1.15	0.63	0.46
MCF7 Spheroid	18.10	18.22	19.03	18.45	0.51
MDA Parental	27.21	28.96	25.66	27.28	1.65
MDA Spheroid	69.63	72.90	69.72	70.75	2.22

% cells stained with ALDH+

	Replicates				
Cell type	1	2	3	Mean	SD
MCF7 Parental	1.18	1.68	2.56	1.81	0.70
MCF7 Spheroid	15.00	15.94	15.20	15.38	0.50
MDA Parental	1.19	1.53	2.21	1.64	0.52
MDA Spheroid	28.20	30.19	29.89	29.43	1.07

Three independent experiments were carried out. (Equivalent to Figure 4.25)

APPENDIX E

The IC50 values of drugs tamoxifen, doxorubicin and cisplatin in parental and spheroid cells (3D and 2D conditions) of MCF-7 breast cancer cell lines

Drug inhibitory concentration, IC50 (μM) - Tamoxifen										
	Replicates									
Cell type	1	2	3	4	5	6	7	8	Mean	SD
MCF-7 Parental	16.8	18.8	21.5	21.5	22.00	22.6	22.3	20.2	20.57	2.25
MCF7 Spheroid (2D)	40.00	40.00	43.50	43.50	50.00	51.00	55.50	57.50	47.63	6.83
MCF7 Spheroid (3D)	53.83	53.83	59.21	61.9	63.25	64.59	75.36	53.82	60.73	7.39

Drug inhibitory concentration, IC50 (μM) - Doxorubicin										
	Replicates									
Cell type	1	2	3	4	5	6	7	8	Mean	SD
MCF-7 Parental	0.51	0.69	0.86	1.20	1.29	1.46	1.28	0.70	1.00	0.37
MCF7 Spheroid (2D)	10.00	10.00	11.00	12.00	13.00	13.00	21.00	16.00	13.25	3.69
MCF7 Spheroid (3D)	13.79	24.13	15.5	17.2	18.9	20.6	21.55	21.54	18.82	3.05

Drug inhibitory concentration, IC50 (μM) -Cisplatin										
	Replicates									
Cell type	1	2	3	4	5	6	7	8	Mean	SD
MCF-7 Parental	1.99	2.16	2.33	2.49	3.33	2.99	2.66	2.33	2.54	0.44
MCF7 Spheroid (2D)	8.00	8.00	12.50	14.00	15.50	16.00	17.50	18.00	13.68	3.92
MCF7 Spheroid (3D)	11.50	12.00	12.50	13.00	14.00	14.00	14.00	15.50	13.61	3.54

Eight independent experiments for each of the drug tested were carried out. (Equivalent to Table 4.3)

APPENDIX F

The IC50 values of drugs tamoxifen, doxorubicin and cisplatin in parental and spheroid cells (3D and 2D conditions) of MDA-MB-231 breast cancer cell lines.

Drug inhibitory concentration, IC50 (μM) -Tamoxifen										
	Replicates									
Cell type	1	2	3	4	5	6	7	8	Mean	SD
MDA Parental	8.61	8.88	9.15	10.77	13.99	10.22	9.20	11.20	10.28	0.27
MDA Spheroid (2D)	33.00	33.50	35.00	36.00	37.00	40.00	44.00	40.50	37.34	3.82
MDA Spheroid (3D)	41.92	38.22	32.07	30.14	41.85	43.09	31.22	43.25	37.72	5.31

Drug inhibitory concentration, IC50 (μM) -Doxorubicin										
	Replicates									
Cell type	1	2	3	4	5	6	7	8	Mean	SD
MDA Parental	0.43	0.51	0.68	0.77	1.03	0.86	0.68	0.50	0.72	1.31
MDA Spheroid (2D)	7.00	7.50	11.00	17.00	11.00	14.00	14.50	15.00	12.13	3.61
MDA Spheroid (3D)	5.17	6.03	8.62	9.48	10.34	12.06	6.89	6.03	8.08	2.44

Drug inhibitory concentration, IC50 (μM) -Cisplatin										
	Replicates									
Cell type	1	2	3	4	5	6	7	8	Mean	SD
MDA Parental	1.33	1.49	1.33	2.33	1.99	1.66	1.99	1.50	1.74	0.38
MDA Spheroid (2D)	2.50	3.00	3.50	4.00	4.00	5.00	5.00	5.00	4.00	0.96
MDA Spheroid (3D)	13.33	13.99	14.66	16.66	19.99	14.99	15.99	15.99	15.67	1.71

Eight independent experiments for each of the drug tested were carried out. (Equivalent to Table 4.3)

APPENDIX G

Cell growth of spheroid and parental cells of MCF-7 and MDA-MB-231 breast cancer cell lines

	Means percentage in reduction in Alamarblue of MCF-7 cell types							
Day	0	3	5	7	10	14	18	21
Parental	0.00	16.87	18.86	22.05	26.68	27.97	21.83	17.73
Spheroid	0.00	20.85	25.24	31.31	50.33	50.35	45.87	45.87
		SD	SD	SD	SD	SD	SD	SD
		1.16	2.78	3.39293	4.17641	1.32	2.11	1.17
		4.52	1.58	2.33	2.18	1.61	6.37	6.37

	Means percentage in reduction in Alamarblue of MDA-MB-231 cell types							
Day	0	3	5	7	10	14	18	21
Parental	0.00	16.27	17.94	21.87	25.90	25.89	27.96	22.96
Spheroid	0.00	17.43	21.61	27.64	42.14	45.23	41.35	43.64
		SD	SD	SD	SD	SD	SD	SD
		0.901	0.741	1.373	2.001	1.761	2.958	2.965
		1.137	0.894	0.620	1.738	1.158	3.276	1.942

Three independent experiments for each of the day were carried out. (Equivalent to Figure 4.36A)

APPENDIX H

Wound healing assay of MCF-7 and MDA-MB-231 spheroid and parental cells

	Percentage of cells migrated to close the wound				
	Replicates				
Cell type	1	2	3	Means	SD
MCF7-Parental	38.523	37.17	28.95	34.88	5.18
MCF-7 Spheroid	49.03	49.43	48.10	48.86	0.68
MDA-Parental	44.36	45.46	59.47	49.77	8.42
MDA-Spheroid	55.38	66.82	62.69	61.42	5.52

Three independent experiments for each of the cell type. (Equivalent to Figure 4.37B and Figure 4.38B)

APPENDIX I

Migration and invasion assay of MCF-7 and MDA-MB-231 spheroid and parental cells

	Percentage of cells migrated cells				
	Replicates				
Cell type	1	2	3	Means	SD
MCF7-Parental	100.00	100.00	100.00	100.00	0.00
MCF-7 Spheroid	279.91	352.51	228.94	287.12	62.10
MDA-Parental	100.00	100.00	100.00	100.00	0.00
MDA-Spheroid	126.93	141.04	122.56	130.17	9.66

	Percentage of cells invaded cells				
	Replicates				
Cell type	1	2	3	Means	SD
MCF7-Parental	100.00	100.00	100.00	100.00	0.00
MCF-7 Spheroid	228.66	257.16	291.31	259.05	31.37
MDA-Parental	100.00	100.00	100.00	100.00	0.00
MDA-Spheroid	122.46	131.12	118.28	123.96	6.55

Three independent experiments were carried out (Equivalent to Figure 4.39 and Figure 4.40)

APPENDIX J

The proportion of cells at different stages in the cell cycle of spheroid and parental MCF-7 and MDA-MB-231 cell line

		Percentage of cells at different cell cycle stage			
Cell type	**Replicates**	**Sub G0/G1**	**G0/G1**	**S**	**G2+M**
MCF7-Parental	1	0.98	90.47	1.17	7.1
	2	1.05	90.43	1.04	7.11
	3	1.04	90.9	1.1	6.54
	Means	1.02	90.60	1.10	6.92
	SD	0.04	0.26	0.07	0.33
MCF-7 Spheroid	1	3.77	77.47	7.29	13.31
	2	4.07	76.62	6.63	14.43
	3	4.29	77.27	7.24	13.31
	Means	4.04	77.12	7.05	13.68
	SD	0.26	0.44	0.37	0.65
Cell type	**Replicates**	**Sub G0/G1**	**G0/G1**	**S**	**G2+M**
MDA-Parental	1	0.35	49	29.05	20.25
	2	0.51	49.18	29.09	19.94
	3	0.69	48.53	29.52	20.04
	Means	0.52	48.90	29.22	20.08
	SD	0.17	0.34	0.26	0.16
MDA-Spheroid	1	2.23	71.97	8.03	17.34
	2	2.46	72.87	7.58	16.65
	3	2.52	72.52	7.96	16.7
	Means	2.40	72.45	7.86	16.90
	SD	0.15	0.45	0.24	0.38

Three independent experiments were carried out (Equivalent to Figure 4.41 and Figure 4.42)

APPENDIX K

Quantification of 5-methylcytosine (5-mC) content of the DNA samples from the parental and spheroids cells of MCF-7 and MDA-MB-231 cell line

	Quantification of 5-mC content of DNA				
	Replicates				
Cell type	1	2	3	Means	SD
MCF7-Parental	1.01	1.03	1.03	1.02	0.01
MCF-7 Spheroid	0.93	0.93	0.92	0.93	0.00
MDA-Parental	0.75	0.76	0.83	0.78	0.04
MDA-Spheroid	0.39	0.39	0.40	0.40	0.01

Three independent experiments were carried out (Equivalent to Figure 4.43B)

APPENDIX L

Fold change of validated miRNAs in MCF7 and MDA spheroid cell types relative to the respective parental controls calculated by comparative ΔΔCt method

(A) MCF-7 spheroid relative to its parental cells						
	RT-qPCR data Fold change (Log 2 transformed)					NGS data Fold change
miRNAs	1	2	3	Mean	SD	Mean
hsa-miR-15b	-3.353	-3.917	-2.789	-3.3530	0.4609	-2.025
hsa-miR-34a	-1.512	-1.615	-1.407	-1.5116	0.0849	-3.968
hsa-miR-148a	-2.152	-2.369	-1.861	-2.1272	0.2081	-3.451
hsa-miR-628	-5.565	-5.591	-5.539	-5.5650	0.0212	-4.580
hsa-miR-196b	-5.305	-5.667	-4.940	-5.3040	0.2970	-2.355
hsa-miR-125b	-4.778	-4.972	-4.583	-4.7775	0.1591	-2.830
hsa-miR-760	-3.254	-3.493	-3.014	-3.2533	0.1956	-10.310
hsa-miR-30c	-5.125	-5.376	-4.874	-5.1250	0.2051	-2.740
hsa-miR-136	3.160	3.871	2.442	3.1575	0.5834	13.960

(B) MDA-MB-231 spheroid relative to its parental cells						
	RT-qPCR data Fold change (Log 2 transformed)					NGS data Fold change
miRNAs	1	2	3	Mean	SD	Mean
hsa-miR-15b	-1.126	-0.681	-0.793	-0.867	0.189	-2.874
hsa-miR-34a	-2.170	-2.383	-1.957	-2.170	0.174	-2.079
hsa-miR-148a	-3.540	-3.763	-3.318	-3.540	0.181	-2.034
hsa-miR-628	-5.133	-4.505	-5.760	-5.133	0.513	-2.140
hsa-miR-196b	-2.462	-2.422	-2.659	-2.514	0.104	-2.321
hsa-miR-204	-5.050	-5.916	-4.184	-5.050	0.707	-5.690
hsa-miR-181a	1.439	1.466	1.456	1.454	0.011	5.810
hsa-miR-205	3.505	3.912	3.098	3.505	0.332	5.763

Means data was taken from a minimum of three biological replicates. (Equivalent to Figure 6.10)