SEQUENCE-SPECIFIC INHIBITION OF MICRORNA-130A GENE BY CRISPR IN NON-SMALL CELL LUNG CANCER CELL LINE

by

NUR AININA BINTI ABDOLLAH

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

March 2019

I humbly dedicate this thesis to my beloved ayah and mama En. Abdollah Ali Musa and Pn. Pauziah Man for their endless love, support and prayers in every step of the way that I take in my life. I would like to thank my siblings, Nur Athirah Nadiah Abdollah, Abdul Hariz Abdollah and Abdul Hafiz Abdollah for always be there for me and encourage me in everything I did.

Thank you.

ACKNOWLEDGEMENT

First and foremost, I would like to offer this endeavour to our God Almighty for the wisdom He bestowed upon me, the strength, peace of mind and good health in order to finish this research. This thesis would not have been possible without an immense help from my supervisor, Dr. Siti Razila Abdul Razak who patiently guide me throughout this long and challenging journey. She is a mentor and motivates me each time I am down either with continuous research work or during writing period. She is willing to share her knowledge and expertise in this study. I would like to give a special thanks to Dr Kumitaa who willingly guided me through the techniques for CRISPR-Cas9 cloning. I would also like to thank my co-supervisors, Dr. Nor Hazwani Ahmad and Professor Dr Narazah Mohd Yusoff for their encouragement and knowledge that they have provided me with. I could not find words to express my gratitude to them and may God Almighty bless them.

This thesis is dedicated to my mom, dad and family members who have always be there with me through my ups and downs and believe that I could do it. Without them, I was nothing. They do not only assist me financially but also extended their support morally and emotionally. I have also received an enormous support from my lab mates throughout my postgraduate study. I am indebted to my lab mates who always listen and gives me words of encouragement and advices. They help me survive all the stress and not letting me give up.

Last but not least, I would like to thank all the staff members and technical staff in AMDI whose services turned my research a success. I am very grateful to have met with the all these people who gives me strength to complete this degree. This accomplishment would not have been possible without them. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

$\Delta\Delta C_t$	Comparative C_T
°C	Degree Celsius
µg	Micro gram
µl	Micro litre
$2^{-\Delta\Delta Ct}$	Fold change
3p	3' end
5p	5' end
A549	Human bronchioloalveolar lung carcinoma
ALK	Anaplastic lymphoma kinase
ATCC	American Type Culture Collection
BACH2	BTB domain and CNC homolog2
Bp	Base pair
Cas	Cluster of CRISPR-related
cDNA	Complementary DNA
CO_2	Carbon dioxide
CRISPR	Clustered regularly interspaced short palindromic repeat
CRMP4	Collapsing response mediator protein 4
crRNA	Small CRISPR RNA
CT	Computer tomography
C_T	Threshold cycle
DAVID	Database for annotations, visualization and integrated discovery
DMSO	Dimethyl sulfoxide
dsDNA	Double stranded DNA

EGF	Epidermal Growth Factor
EGFR	Epidermal growth factor receptor
EXP5	Exportin 5
FBS	Fetal bovine serum
FOSL1	Fos-related antigen 1
GFP	Green fluorescent protein
HEK293T	Human embryonic kidney
Hr	Hour
IGF-1	Inhibits insulin-like growth factor 1
IHME	Institute for Health Metrics and Evaluation
Kasumi-1	Human acute myeloid leukemia cell line
Kb	Kilo base pair
KEGG	Kyoto Encyclopaedia of Genes and Genomes
lentiCRISPR-3p	lentiCRISPRv2 cloned with miR130a-3p sgRNA
lentiCRISPR-5p	lentiCRISPRv2 cloned with miR130a-5p sgRNA
lentiCRISPR-SL	lentiCRISPRv2 cloned with miR130a-SL sgRNA
MCF10A	Human mammary epithelial cell line
MCF7	Human breast adenocarcinoma
MET	Mesenchymal-epithelial transition factor
Mg	Mili gram
MGB	Minor groove binder
Min	Minute
miRNA	MicroRNA
mL	Mili litre
Mm	Mili meter
NCBI	National Centre for Biotechnology Information

NF-kB	Nuclear factor-*kB*
NFQ	Nonflorescent quencher
NSCLC	Non-small cell lung cancer
Nt	Nucleotide
oncomiR	Oncogene miRNA
PAM	Protospacer adjacent motif
PBS	Phosphatase buffered saline
PCR	Polymerase chain reaction
PD-1	Programmed cell death 1
pre-miRNA	Precursor-miRNA
pri-miRNA	Primary-miRNAs
PTEN	Phosphatase and tensin homolog on chromosome 10
qRT-PCR	Quantitative real time-polymerase chain reaction
RAB5A	Ras-related protein Rab-5A
RISC	RNA-induced silencing complex
RNU6B	Small nucleolar U6
Rpm	Revolutions per minute
RT	Reverse transcription
RT	Room temperature
RUNX3	Runt-related transcription factor 3
SCLC	Small-cell lung cancer
Sec	Seconds
sgRNA	Single guide RNA
SL	Stem loop
snoRNAs	Small nucleolar RNAs
SVM	Support vector machine

TALENS	Transcription activator-like effector nuclea ses
TKI	Tyrosine kinase inhibitors
TNF-α	Tumour necrosis factor-alpha
tracrRNA	Trans-activating crRNA
UTR	Untranslated region
WHO	World Health Organisation
ZFN	Zinc-finger nucleases

PERENCATAN SPESIFIK URUTAN GEN RNA MIKRO-130A MENGGUNAKAN CRISPR DALAM TITISAN SEL KANSER PARU-PARU SEL BUKAN KECIL

ABSTRAK

Kanser paru-paru sel bukan kecil (NSCLC) adalah salah satu punca utama kematian di dunia dan insiden telah meningkat selama bertahun-tahun. Terapi semasa bergantung kepada radiasi dan kemoterapi, yang pada masa yang panjang akan memaparkan rintangan radiografi dan kemoterapi di lokasi kanser. Oleh itu, untuk meningkatkan kecekapan dan kekhususan dalam terapi kanser, strategi terapi yang disasarkan berdasarkan RNA mikro (miRNA) sedang dibangunkan. Beberapa kajian telah melaporkan bahawa miRNA memainkan banyak peranan dalam pelbagai tahap karsinogenesis kanser, dan boleh berfungsi sebagai alat terapeutik yang berpotensi untuk pengurusan kanser. Oleh itu, kajian ini bertujuan untuk mengkaji peranan miR130a dalam NSCLC dengan memanipulasi ungkapan miRNA menggunakan sistem ulangan palindromik (CRISPR)-Cas9 secara berkala. Bagi mencapai tujuan itu, 20 bp urutan RNA berpandu tunggal yang disasarkan pada gelung batang, 3 'dan 5' hujung miR130a direka dan diklonkan ke dalam plasmid lentiCRISPRv2. Klon positif telah disahkan oleh penjujukan dan transjangkitan masuk ke sel-sel A549 menggunakan Lipofectamine® 3000 sebelum sel-sel transjangkitan stabil dipilih oleh puromycin. Keseluruhan RNA diekstrak dan tahap miR130a-3p dan miR130a-5p matang dikira menggunakan qRT-PCR menggunakan TaqMan Universal Master Mix (Applied Biosystems). Keputusan menunjukkan bahawa ekspresi miR130a-3p dan 5p secara signifikan dikurangkan kepada dua dan tiga kali ganda. Daya tahan sel telah dianalisis menggunakan ujian MTS dan ditunjukkan bahawa penurunan ekspresi miR130a-3p dan 5p secara signifikan merencatkan kebolehidupan sel tahan sel pada 24 dan 48 jam. Kesan proliferasi jangka

panjang oleh ujian pembentukan koloni dilakukan dan hasil menunjukkan penurunan ketara dalam proliferasi sel seperti yang ditunjukkan oleh penurunan jumlah dan saiz kolon. Oleh itu, keputusan menunjukkan bahawa sistem CRISPR-Cas9 boleh digunakan untuk menindas ungkapan miR130a-3p dan 5p matang ke atas sel A549. MiRNA bertindak sebagai onkogen ke atas sel-sel A549 disebabkan oleh penurunan ekspresi tersebut telah menunjukkan penurunan kebolehidupan sel-sel A549 dengan signifikan. Untuk meramal fungsi peraturan miR130a, alat analisa bioinformatik gabungan dilakukan untuk mengenal pasti miR130a-3p dan 5p gen sasaran yang berpotensi. Analisis telah mengenal pasti sembilan gen sasaran yang diramalkan untuk miR130a-3p (RAPGEF4, SOS2, NRP1, RPS6KB1, MET, IL15, ACVR1, RYR2 and ITPR1), sepuluh untuk miR130a-5p (BCL11A, SPOPL, NLK, PPARGC1A, POU4F2, CPEB4, ST18, RSBN1L, ELF5 and ARID4B). Walau bagaimanapun, kajian-kajian hiliran yang lebih lanjut antara gen sasaran yang diramalkan dan miR130a adalah perlu untuk lebih memahami interaksi mereka. Kesimpulannya, kajian ini menunjukkan potensi terapeutik miR130a sebagai terapi sasaran peribadi untuk merawat NSCLC.

SEQUENCE-SPECIFIC INHIBITION OF MICRO RNA-130A GENE BY CRISPR IN NON-SMALL CELL LUNG CANCER CELL LINE

ABSTRACT

Non-small cell lung cancer (NSCLC) is one of the main causes of mortality globally and the incidence has been rising over the years. Current therapies rely on radiation and chemotherapy, which over time exhibited radio-resistance and chemoresistance at the cancer site. Hence, to improve the efficiency and specificity in the cancer therapy, targeted therapeutic strategies based on microRNA (miRNA) are currently being developed. Several studies have reported that miRNA play roles in multi-steps carcinogenesis of cancers, and may serve as potential therapeutic tools for cancer management. Therefore, this study aims to investigate the roles of miR130a in NSCLC by manipulating the expression of miRNA using the clustered regularly interspaced short palindromic repeat (CRISPR)-Cas9 system. To achieve the aim, 20 bp single-guided RNA sequences targeted at stem loop, 3' and 5' end of miR130a were designed and cloned into lentiCRISPRv2 plasmid. The positive clones were confirmed by sequencing and transfected to A549 cell line using Lipofectamine® 3000 before the stably transfected cells were selected by puromycin. Total RNA was extracted and the mature level of miR130a-3p and miR130a-5p were quantified using qRT-PCR using TaqMan Universal Master Mix (Applied Biosystems). Result showed that the expression of miR130a-3p and 5p were significantly downregulated to two and three-fold respectively. Cell viability was analysed using MTS assay and it was shown that downregulation of miR130a-3p and 5p significantly inhibit cell viability at 24 and 48 hours. Long-term proliferation effect by colony formation assay was performed and result showed significant decrease in cell proliferation as exhibited by decrease number and size of colonies. Thus, result suggested that the CRISPR/Cas9 system can be used to suppress the expression of mature miR130a-3p and 5p in A549 cell line. The miRNAs acted as oncogene in A549 cell line as

downregulation of them significantly suppressed the viability of A549 cell line. To predict the regulatory function of miR130a, combinatorial bioinformatics analyses tool was performed to identify potential target genes of miR130a-3p and 5p. The analyses have identified nine predicted target genes for miR130a-3p (RAPGEF4, SOS2, NRP1, RPS6KB1, MET, IL15, ACVR1, RYR2 and ITPR1), and ten for miR130a-5p (BCL11A, SPOPL, NLK, PPARGC1A, POU4F2, CPEB4, ST18, RSBN1L, ELF5 and ARID4B). However, further downstream studies between these predicted target genes and miR130a are necessary to further understand their interaction. In conclusion, this study has shown the therapeutic potential of miR130a as personalised targeted therapy for the treatment of NSCLC.

CHAPTER 1

INTRODUCTION

1.1 Background of the study

Around 42 million people in the world has been estimated to suffer from any forms of cancer. In 2016, the highest prevalence of cancer are breast cancer with 8 million, colon and rectum with 6.3 million, prostate with 5.7 million and trachea, bronchus and lung cancer with over 2.8 million (Global Burden of Disease Collaborative Network, 2017). Although lung cancer does not rank as the highest in prevalence, it causes the highest mortality globally with 1.71 million which is more than double the number of the second ranked, stomach cancer (0.83 million) (Global Burden of Disease Collaborative Network, 2017). Besides that, World Health Organisation (WHO) reported that trachea, bronchus and lung cancer ranked 6th in the top ten cause of death in the world with 23 crude death per 100 000 population (World Health Organisation, 2018b).

Lung cancer is malignant tumour which starts in bronchial epithelium, bronchioles, alveoli, and bronchial glands (Zamay et al., 2017). WHO classified lung cancer into two main types which are small-cell lung cancer (SCLC) and non-small lung cancer (NSCLC). NSCLC attributed to 85% of lung cancer cases (Wood et al., 2018). Approximately only 13 – 15% of lung cancer patients survive within five years after diagnosis (Zamay et al., 2017). Detection at later stage causes them to spread from lung to central nervous system in 54% cases (Gaspar et al., 1997, Cho et al., 2005). Difficult for early detection and resistance towards available radiotherapy and chemotherapy contributes for high morbidity and mortality in lung cancer cases (Sato et al., 2007).

Most lung cancer cases were detected at an advanced stage when it has already metastasised, and only when symptoms such as coughing, coughing up blood, chest pains

and shortness of breath developed (Hirsch et al., 2017). Patient with earlier stage is suggested to undergo surgery (Hirsch et al., 2017). Radiotherapy and chemotherapy were carried out for patient with medical contraindications for surgery, for patient who refuses surgery and for advanced stage lung cancer patient (Hirsch et al., 2017). However, current radiotherapy and chemotherapy have been shown to exhibit resistance. Thus, targeted therapy for specific molecular genomics alteration may be a better therapy option for patient. To support, a recent study by Merck revealed that in comparison to chemotherapy, pembrolizumab, programmed cell death 1 (PD-1) inhibitor increases survival benefit for patients with PD-L1-expressing tumours in NSCLC (Kenilworth, 2016). In addition, a study suggested an effective targeted agent in conjunction with radiotherapy and chemotherapy for locally advanced NSCLC patient is needed to increase the survival rate of patient (Hirsch et al., 2017). Hence, targeted therapy on molecular genomic alteration will improve the efficiency and specificity in lung cancer therapy.

The current therapy rarely cure the cancer as the survival rate within 5 years is only 15% (Sato et al., 2007). Thus, there is an important need for new tools which use the knowledge about specific molecular targets and the pathogenesis of lung cancer (Fong et al., 1999). These tools need to address on how the tumour cell is regulated by oncogenes and tumour suppressor genes by causing uncontrollable proliferation and metastasis via disruption of important cell-cycle regulators and cascades of signal transduction (Fong et al., 1999). Currently, only two genetic abnormalities namely epidermal growth factor receptor (EGFR) mutation and anaplastic lymphoma kinase (ALK) gene rearrangements were developed with targeted therapies (Hirsch et al., 2017). Therefore, specific and highly sensitive reliable biomarkers and personalised molecular targeted therapies are important to increase specificity of treatment as well as to increase the survival rate of the advanced stage lung cancer patients.

In order to investigate the target molecular activity of genes, selection of precise and efficient gene editing technique is important. One of the genome editing technology, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system is increasingly being employed to silence protein-coding genes in several model organisms (Ran et al., 2013). CRISPRs are DNA loci with short repetitions of base sequences (Marraffini and Sontheimer, 2010). Cas9 is nuclease guided by small RNAs introduces dsDNA cleavage at targeted genomic locus (Ran et al., 2013). Cas9 reduce the off-target effect as it is highly specific, efficient and well-suited for high throughput and multiple gene editing for different cell types and organisms (Ran et al., 2013, Mali et al., 2013). However, the use of this CRISPR-Cas9 technologies need to be further investigated at both pre-clinical and clinical level to study the safety, specificity and efficiency of the system in lung cancer targeted therapy.

In this study CRISPR-Cas9 systems is used to increase the understanding of regulation and function of miR130a in NSCLC cell line. MicroRNAs (miRNAs) are short stranded noncoding RNA of ~22 nucleotides that participates in various human biological activities such as regulation of apoptosis, cell proliferation, cellular differentiation, development and immune responses (He and Hannon, 2004). In cancer cells, several studies have shown that miRNAs have an ability to distinguished malignant cancer cells from normal and play a role in the carcinogenesis development in cancer patients giving high hope to use miRNAs as biomarker for diagnosis, prognosis and treatment of cancer (Li et al., 2008, Wang et al., 2014, Nurul-Syakima et al., 2011, Chen et al., 2009). The miR-130a has been reported to be upregulated in several types of cancer, such as colon cancer, chronic myeloid leukaemia and hepatocellular carcinoma indicating the important roles of miR130a in cancer development, progression and metastasis (Chen et al., 2010, Wang et al., 2010, Suresh et al., 2011, Zhou et al., 2012). Thus, the role of miR130a need

to be extensively studied at *in vitro* level first in order to understand the regulatory function of miR130a in lung cancer.

Association between miRNAs and target genes can be predicted using several miRNA prediction databases. The bioinformatics platforms have been used to analyse large amount of data and to perform comparative analysis of differentially expressed genes. Comparative studies of the earlier miRNA target prediction programs proved that there were no program that constantly better than all others (Rajewsky, 2006, Sethupathy et al., 2006). The platforms may be used together to complement each other in order to achieve the objective of predicting the target genes of miR130a. Hence, in this study, three bioinformatics platforms namely miRWalk, the Database for annotations, visualization and integrated discovery (DAVID) Gene Functional Classification Tool and miRanda-miRSVR analysis tools were used to identify possible interaction between miR130a and its target, which eventually will be used to understand the mechanisms of interactions in NSCLC.

1.2 Problem statement

NSCLC has been attributed to largest cases to lung cancer patients. Thus, development of specific and highly sensitive new targeted therapies is important to improve current therapies as well as to increase the survival rate of the advanced stage lung cancer patients. miR130a has been shown to play important role in carcinogenesis of lung cancer. However, the role of miR130a is not well documented in NCSLC. One of the ways to study the function of miR130a in NSCLC is by gene manipulation by silencing the expression of miR130a. Gene manipulation can be performed by several means, and that includes the genome editing technology, CRISPR-Cas9 system. However, the efficiency and specificity of CRISPR-Cas9 system in silencing the expression of miRNA is poorly studied. Hence, in this study, CRISPR-Cas9 system will be implemented to silence the expression of miR130a in NSCLC cell line. This CRISPR-Cas9-miRNA system may assist in understanding the multistep mechanisms of miR-130a regulatory activities in NSCLC. Besides that, in order to identify the target gene of miR130a in pathways involved, three bioinformatics platforms namely miRWalk, DAVID gene functional classification tool, miRanda-miRSVR analysis tool were employed for gene prediction analysis of miR130a-3p and 5p and STRING database for protein-protein interaction analysis between predicted target genes.

1.3 Research objective

The ultimate objective of this study is to investigate the role of miR130a by knocking-down the expression of miR130a using CRISPR-Cas9 system, and to predict associations between miR130a and target genes by bioinformatics analyses.

1.3.1 Specific objectives:

i. To design single guide RNA targeting at the miR130a gene and to clone in the CRISPR Nuclease Vector.

ii. To establish stable miR130a suppression in A549 cancer cell line.

iii. To investigate the effects of miR130a suppression on viability and proliferation of A549 cell line.

iv. To identify potential target genes of miR130a and to predict their associations by bioinformatics analysis

1.4 Hypotheses

The CRISPR/Cas9 system could be used to stably silence the expression of miR130a. Silencing of miR130a affect the oncogenic properties of the A549 cell line particularly in the regulation of cell viability and cell proliferation. The bioinformatics analyses will identify potential target genes of miR130a, and predict the association between miR130a and its target genes in NSCLC.

CHAPTER 2

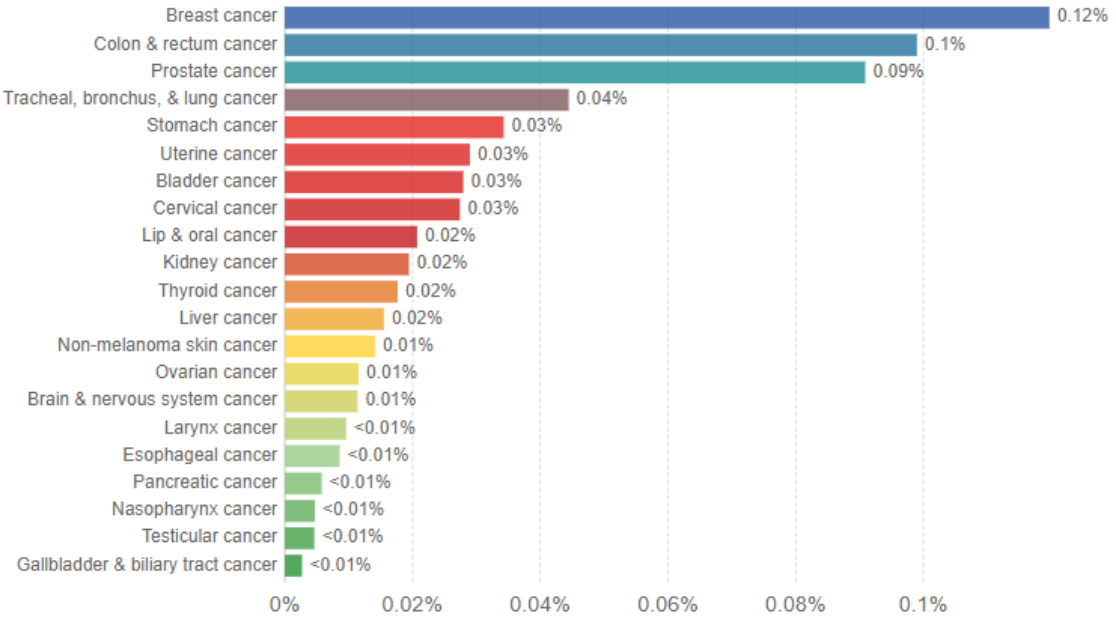
LITERATURE REVIEW

2.1 Cancer

Cancer is a main global public health problem. According to the World Health Organisation (WHO), 8.8 million deaths in 2015 is caused by cancer (World Health Organisation, 2018a). One in every six deaths worldwide is caused by cancer, making it the second highest cause of death in the world (Naghavi et al., 2017). Similarly, cancer is also a major health problem and the leading cause of death in Malaysia. Recently, Department of Statistics Malaysia reported that cancer ranked as the top five principal cause of death in 2016 (Department of Statistics Malaysia, 2017). Cancer is a large group of diseases which caused by a rapid growth of abnormal cells that have potential to spread to other organs via blood or lymph systems (World Health Organisation, 2018a). It can start almost anywhere in the body due to the interaction between genetic factors and external agents. External agents can be categorised into physical carcinogens (e.g. ultraviolet and ionising radiation), chemical carcinogens (e.g. asbestos, components in tobacco smoke) and biological carcinogens (e.g. infections from certain viruses, bacteria or parasites) (World Health Organisation, 2018a).

There are over 200 types of cancers identified so far (World Health Organisation, 2018a). National Cancer Institute has categorised them under general names which are carcinoma, sarcoma, leukaemia, lymphoma and myeloma and central nervous system cancer (National Cancer Institute, 2015). Carcinoma is cancer that starts in the skin or in tissue that cover internal organs. Sarcoma on the other hand is cancer which starts in bone, cartilage, fat, muscle, blood vessels or other connective or supportive tissue. Cancer of blood is known as leukaemia while lymphoma and myeloma are cancers which starts in cells of the immune system. Lastly, cancer which begins in the tissue of the brain or spinal

cord such as brain and spinal cord tumours are known as central nervous system cancers (National Cancer Institute, 2015). Statistics from the Institute for Health Metrics and Evaluation (IHME) and Global Burden of Disease showed that the four leading types of cancer are breast cancer, colon and rectum cancer, prostate cancer, and tracheal, bronchus and lung cancer (**Figure 2.1**).



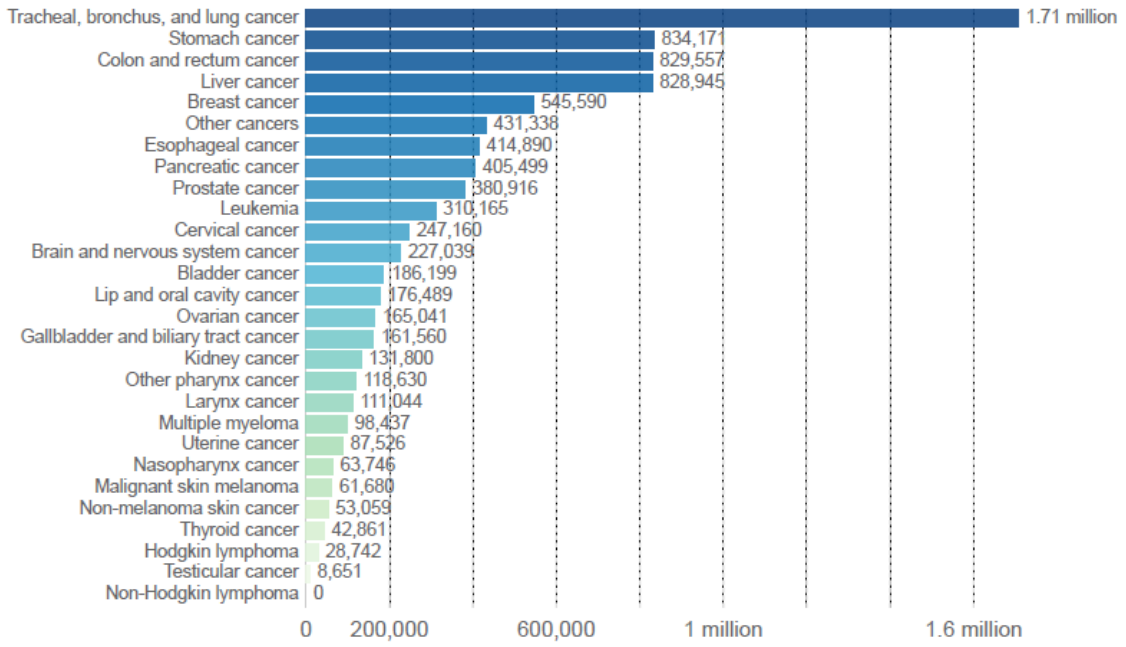
Source: IHME and Global Burden of Disease (2017)

Figure 2.1: World cancer incidence in 2016 among men and women in all age groups (Global Burden of Disease Collaborative Network, 2017).

2.2 Lung cancer

Lung cancer is the major cause of mortality and morbidity globally with 1.71 million deaths in 2016 (**Figure 2.2**) (Naghavi et al., 2017). The incidence of lung cancer deaths is expected to increase until 2030 unless efforts for worldwide tobacco control are prominently increased (Wood et al., 2018). In 2014, WHO reported that lung cancer recorded for 19.1 deaths per 100,000 population in Malaysia or 4,088 deaths per year (3.22% of all deaths) (World Health Organisation, 2018a). In addition, malignant neoplasm of trachea, bronchus and lung was also ranked as top five leading cause of death

in Malaysia in 2016 which showed that lung cancer is the most common cause of deaths in Malaysia (Department of Statistics Malaysia, 2017).



Source: IHME and Global Burden of Disease (2017)

Figure 2.2: World mortality numbers in different types of cancer in 2016 among men and women in all age groups (Global Burden of Disease Collaborative Network, 2017).

2.2.1 Types of lung cancer

Lung cancer is an uncontrolled growth of abnormal cell in tissues of the lung. It can be metastasised into nearby tissues or other parts of body. Most cancers which start in lung (primary lung cancers) are categorised as carcinomas (National Cancer Institute, 2015). WHO divided lung cancer based on its biology, therapy and prognosis. Two major classes of lung cancer are small cell lung cancers (SCLC) and NSCLC, and about more than 85% of lung cancer cases were attributed to NSCLC (American Cancer Society, 2017). There are three different types of NSCLC namely adenocarcinoma, squamous cell carcinomas and large cell carcinomas which are segregated based on the type of cells found in the tumour (Wood et al., 2018). Adenocarcinoma is the most common type and it starts in the

mucus making gland cells in the lining of airways (American Cancer Society, 2017). Squamous cell carcinoma on the other hand starts in even and smooth cells surface which cover the external layer of airways and have a tendency to grow near the centre of lung (American Cancer Society, 2017). Cancer cells which is large and round when observed under microscope is categorised as large cell carcinomas (American Cancer Society, 2017). In lung cancer, lack of reliable markers, sensitive tools and poor prognosis contribute to poor early diagnosis and patient care (Ludwig and Weinstein, 2005).

2.2.2 Diagnostic tests to screen for lung cancer

Cigarette smoking is the main risk factor for development of lung cancer as nearly half of reported cases are diagnosed in former smoker (Wood et al., 2018, Sato et al., 2007). Therefore, identifying former smoker has been a crucial part in determining the highest risk individual to develop lung cancer (Sato et al., 2007). It is important to note that most cases were often detected by accident and first noticed at advanced stages when the cancers have already spread widely (Sato et al., 2007, Zamay et al., 2017). Efficient early detection method is important as it may help in earlier treatment of patient. Sputum cytology, chest x-ray and computer tomography (CT) scan has been used to screen for lung cancer (American Cancer Society, 2017). Currently, new diagnostic tests such as polymerase chain reaction (PCR), next-gene sequencing and florescent *in-situ* hybridisation have been employed for a better and precise screening for specific lung cancer genomic abnormalities (Hirsch et al., 2017). Previous studies have reported that the development of lung cancer is from a multi-stage carcinogenesis of normal epithelial cells which are caused by genetic and epigenetic abnormalities, usually coupled with cigarette smoking (Zöchbauer-Müller et al., 2002, Sato et al., 2007). Currently, there are no one reliable therapy for different stages of lung cancer (Sato et al., 2007). Thus, novel, highly sensitive and specific molecular gene target are crucial for therapy in different stages of lung cancer patient.

2.3 Current therapy for lung cancer

The recommended treatment for patients with NSCLC (stage I and II) are surgery (Cao et al., 2012). Recently, molecular targeted therapies have been a great help for the treatment of advanced lung cancer patient, specifically the one with specific genomic abnormalities (Hirsch et al., 2017). A study stated that about 69% of patient with advanced lung cancer stage may have a potentially genomic molecular target (Tsao et al., 2016). **Figure 2.3** shows the targeted therapies for specific molecular features of genomic abnormalities (Hirsch et al., 2017). The most common Epidermal Growth Factor Receptor (EGFR) mutations in NSCLC patients are deletions in exon 19 and mutations in exon 21 which results in the activation of the tyrosine kinase domain. The approved drugs which act as tyrosine kinase inhibitors (TKI) are erlotinib, gefitinib and afatinib (Langer, 2013). Another genomic aberration with approved drugs is anaplastic lyphoma kinase (ALK) gene rearrangements. There are about 2 to 7 percent of NSCLC patients with the ALK rearrangements (Douillard et al., 2014). They have some comparable clinical characteristics to patients with EGFR mutations but resistant to EGFR TKIs (Ettinger et al., 2017). The approved drugs for patients with ALK rearrangements are crizotinib, which is an inhibitor for ALK, receptor tyrosine kinase 1 (ROS1) and some mesenchymal-epithelial transition factor (MET) tyrosine kinase. Patients whose disease progress with crizotinib will be given alectinib as a second option (Kim et al., 2016, Cheng and R Ott, 2010). Certinib also act as TKI of ALK which also inhibits insulin-like growth factor 1 (IGF-1) receptor (Ettinger et al., 2017). Currently, there are no approved drugs for the other genomic mutations listed in **Table 2.1** besides EGFR mutations and ALK rearrangements. As current standard drugs only give a modest survival benefits, a personalised drugs are better suited to molecular phenotypes of each individual which may increase the survival rate of patient with lung cancer (Sato et al., 2007).

		Molecular features	Diagnostic tests	Approved drugs
Mutations	EGFR	>90% exon 19 deletions and exon 21 Activating mutations, exclusive of other driver aberrations; Thr790Met mutations are present in about half of the cases after tyrosine kinase inhibitor resistance	PCR, next-generation Sequencing (mandatory in non-SCC or non-smokers with NSCLC)	Gefitinib, erlotinib, afatinib, osimertinib (Thr790Met, resistant disease)
	KRAS	Codon 12 (>90%), codon 13 (<10%); exclusive of other driver aberrations	PCR, Sanger, next-generation sequencing	None
	BRAF	Activating mutations of the tyrosine kinase domain; about half are Val600Glu mutations	PCR, Sanger, next-generation sequencing	None
	HER2	Activating exon 20 insertions; exclusive of other driver aberrations	PCR, Sanger, next-generation sequencing	None
	MET	Exon 14 skipping mutations; some overlap with MET amplification and other drivers	PCR, Sanger, next-generation sequencing	None
Rearrangements	ALK	Fusions of partner gene with exon 20 of *ALK*; partners in the 27 known fusion variants are *EML4*, *KIF5B*, *TFG*, and *KLC1*; exclusive of other driver aberrations	FISH, immunohistochemistry (mandatory in non-SCC or nonsmokers with NSCLC)	Crizotinib, ceritinib, alectinib (crizotinibresistant)
	ROS1	Nine fusion proteins described with the *FIG*, *SCL34A2*, *TPM3*, *SDC4*, *EZR*, *LRIG3*, *KDELR2*, and *CCDC6* genes; exclusive of other driver aberrations	FISH, immunohistochemistry (recommended in EGFR and ALK with adenocarcinoma)	None
	RET	Fusions described with four partner genes: *KIF5B*, *CCDC6*, *NCOA4*, and *TRIM33*; exclusive of another driver Aberrations	FISH	None
	NTRK	Fusions of *NTRK1* and *NTRK2* occur with a range of partners; *NTRK3* fusions are rare	FISH	None
Copy number gain	MET	Several definitions; preferable: MET:CEP7 ratio >5; might be associated with other driver anomalies if definition criteria not strict	FISH	None
PCR: Polymerase Chain Reaction, FISH: Fluorescent *In-Situ* Hybridisation				

Table 2.1: Diagnostic tests and molecular targeted therapies based on genetic alterations of lung cancer (Hirsch et al., 2017).

2.4 Gene therapy: CRISPR-Cas9 system

Zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENS), and CRISPR-Cas9 system are new methods that have been introduced to manipulate gene expression in an *in vitro* and *in vivo* models (Gaj, 2013). However, CRISPR-Cas9 system is preferable as it is a simple, flexible and more cost-efficient method with an ability to produce a stable knockout in different types of cells (Chang et al., 2016). CRISPR-Cas9 has been shown to be successfully used in gene knockout including in human cell lines, mice, zebrafish, *Drosophila* and *Caenorhabditis elegans* (Hwang et al., 2013, Mali et al., 2013, Pépin et al., 2014, Ran et al., 2013). The establishment of this new genetic manipulation may lead to a successful gene therapy in the future. CRISPR or clustered, regularly interspaced, short palindromic repeats - associated (Cas) systems is used in bacteria to silence foreign nucleic acids (Hwang et al., 2013). CRISPRs are DNA loci with short repetitions of base sequences (Marraffini and Sontheimer, 2010). Short segments of "spacer DNA" are present in between each repetitions from previous exposures to a virus (Marraffini and Sontheimer, 2010).

There are three types of CRISPR system (I – III), each containing different subtypes. Each system comprises a cluster of CRISPR-related (Cas) genes, noncoding RNAs and a distinctive direct repeats of arrays (Ran et al., 2013). The process and transcription of the repeat-spacer array forms the accumulation of small CRISPR RNAs (crRNAs) containing a full or partial sequence (Bikard and Marraffini, 2013). crRNAs play a role to guide the Cas9 ribonucleoprotein complex to the protospacers, a short variable sequence on the exogenous DNA targets (Bikard and Marraffini, 2013, Ran et al., 2013, Gottesfeld et al., 1997). Within the DNA targets sequences, each protospacer always accompanied with a protospacer adjacent motif (PAM), which is different in differing CRISPR-Cas9 systems (Marraffini and Sontheimer, 2010, Bikard and Marraffini, 2013, Barrangou et al., 2007, Brouns et al., 2008). Type II CRISPR system consists of nuclease

Cas9, the trans-activating crRNA (tracrRNA) and a small RNA antisense to the repeat sequence. Each crRNA unit contains 20-nucleotide (nt) guide sequence and a partial direct repeat which guide Cas9 to a 20-base pair (bp) DNA target (Ran et al., 2013). The crRNA and tracrRNA can be merged together forming a chimeric, single guide RNA (sgRNA) (Ran et al., 2013). Different Cas9 orthologs may have different PAM site (Makarova et al., 2011). Cas9 scans region which is complementary between the sgRNA to the double stranded DNA (dsDNA) genome of the invader (Bikard and Marraffini, 2013). Cas9 then introduces dsDNA cleavage using two independent nuclease domains (RuvC and HNH) on each of the DNA strand (Bikard and Marraffini, 2013). As the CRISPR system used in this study was derived from *Streptococcus pyogenes*, the target DNA must immediately precede a 5′-NGG-3′ PAM site for cleavage to occur (Jinek et al., 2012, Bikard and Marraffini, 2013) (**Figure 2.3**).

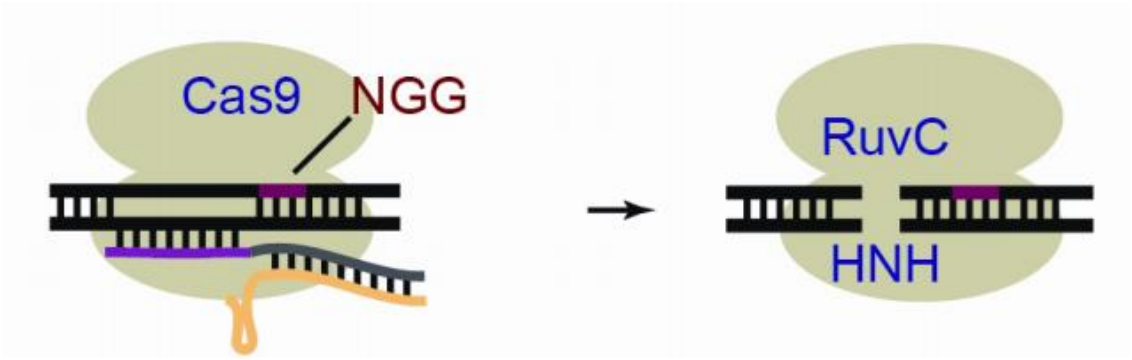


Figure 2.3: Cas9 causing dsDNA breaks on both strands of the genome targeting based on sgRNA sequence in type II CRISPR-Cas9 system with the presence of 5′-NGG-3′ PAM site on targeted dsDNA (Bikard and Marraffini, 2013).

Cas9 decrease the off-target effect as it is highly specific, efficient and well-suited for high throughput and multiple gene editing for different cell types and organisms (Ran et al., 2013, Mali et al., 2013). This double-nicking strategy using CRISPR-Cas9 approach

has successfully generate stable knockout cell line models in human (Pépin et al., 2014, Mali et al., 2013). Recently, CRISPR-Cas9 system has been effectively used for knockout AKRIBI gene expression in human endometrial cell lines (Pépin et al., 2014). Previously, Mali et al., (2013) has also shown the use of CRISPR-Cas9 as a tool in gene editing in human cell line by engineering the CRISPR system of type II bacterial to function with custom gRNA in human cells (Mali et al., 2013). CRISPR-Cas9 systems will ease the efficacy of human editing. Thus, in this study CRISPR-Cas9 systems is used to increase the understanding of regulation and function of miR130a in NSCLC cell line. Consequently, these models can be carried out in human tissues and complement animal studies for future studies. Thus, the application of CRISPR-Cas9 system in gene therapy need to be extensively studied at *in vitro* and *in vivo* level first which can eventually benefit not only to treat cancer but for other diseases as well.

2.5 miRNA as a target gene for therapy

Several studies reported that microRNA (miRNA) expression profiles and specific miRNAs were able to distinguish between better-prognosis and worse-prognosis of lung cancers (Torre et al., 2015). As miRNAs circulates in the body fluids, they may serve as promising biomarkers for early diagnosis and predicting prognosis of lung cancers (Torre et al., 2015). Thus, in this study, miRNA was used as target gene for CRISPR-Cas9 system. miRNAs are short stranded noncoding RNA of ~22 nucleotides that participates in various human biological activities such as regulation of apoptosis, cell proliferation, cellular differentiation, development and immune responses (He and Hannon, 2004). miRNA were firstly discovered because of their roles in controlling the timing of larvae development in *Caenorhabditis elegans* (Hurst et al., 2009). Later, it was then discovered in plant and mammalian cells (Filipowicz et al., 2008). To date, 2588 of mature human miRNAs have been identified (Kozomara and Griffiths-Jones, 2014) and the expression pattern of these miRNAs are regulated both spatially and temporally in the cells (Jeong

et al., 2014, Ziu et al., 2014). These miRNAs can be found in tissue as well as in the circulating peripheral blood, and predicted to control various metabolic pathways, regulate differentiation of stem cells and maintain biological homeostasis (Liu et al., 2013, Vasanthan et al., 2015, Yong et al., 2015). In cancer cells, several studies have shown that some miRNAs have an ability to distinguish malignant cancer cells from normal and play a role in the carcinogenesis development in cancer patients giving high hope to use miRNAs as therapeutic target for diagnosis, prognosis and treatment of cancer (Li et al., 2008, Wang et al., 2014, Nurul-Syakima et al., 2011, Chen et al., 2009).

2.5.1 Silencing of miRNA *in vitro*

miRNA loss-of-function studies have been employed to manipulate the expression of miRNA. Three general approaches in silencing miRNAs are genetic knockouts, miRNA sponges and antisense oligonucleotides (**Table 2.2**) (Meister et al., 2004, Krützfeldt et al., 2005, Ørom and Lund, 2010). Generation of miRNA gene knockouts has extensively revealed on the functions of miRNAs in *C. elegans* and *Drosophila* and more than 476 miRNAs in mouse (Williams et al., 2009, van Rooij et al., 2009, Sokol and Ambros, 2005, Abbott et al., 2005, Miska et al., 2007, Johnnidis et al., 2008, Ventura et al., 2008, Van Rooij et al., 2008, Mu et al., 2009). miRNA sponge consists of a numbers of target sites which complement to target miRNA (Ebert and Sharp, 2010). It then specifically silences the activity of a family of miRNAs with a common seed both in transient and long-term inhibition (Stenvang et al., 2012). AntimiR oligonucleotides is the first modified antisense oligonucleotides that has been demonstrated to work in mammals (Broderick and Zamore, 2011). AntimiR oligonucleotides composed of 2′ -O-methyl-modified ribose sugars, terminal phosphorothioates and at the 3′end cholesterol group which assist in the delivery of antimiR to cells (Broderick and Zamore, 2011). Later, locked nucleic acid (LNA), an alternative modified technology has been introduced into the 2′ -O-methyl- modified ribose sugars structure (Broderick and Zamore, 2011). LNA addition to the structure

increases the binding affinity to target miRNA and are more resistant to degradation (Broderick and Zamore, 2011). However, miRNA sponges and antimir oligonucleotides are difficult to synthesis, costly and causes toxic side-effects which limit on the application (Kluiver et al., 2012, Garzon et al., 2010). Thus, in this study, CRISPR-Cas9 system was employed to silence on the expression on miRNA as it is a simple, flexible and more cost-efficient method in generating a stable transfection.

Table 2.2: Strategies to silence miRNA activity (Stenvang et al., 2012).

Technology	Characteristics	*In vitro*
Genetic knockout	Constitutive or conditional	Primary cells
miRNA sponges	Transient to long-term	Transfection or viral delivery
antimiR oligonucleotides	Transient (in vitro) to long-lasting inhibition	Transfection or unassisted uptake

2.52 Biogenesis of miRNA

The miRNAs post-transcriptionally regulate the expression of protein-coding genes by binding to the specific binding sites in their 3' untranslated region (UTR) and cause either mRNA degradation or translational termination (Bartel, 2004). There is a scarce knowledge about the transcription regulation of primary-miRNAs (pri-miRNA). However, a study showed that certain pri-miRNA is sited in the introns of host genes for both protein-coding genes and non-coding genes (Lagos-Quintana et al., 2003). Thus, it may be transcriptionally regulated by their host-gene promoters (Lagos-Quintana et al., 2003). In nucleus, pri-miRNA is transcribed by RNA polymerase II (Ha and Kim, 2014). As shown in **Figure 2.4 (a)**, the pri-miRNA is made up of a stem of 35-35 bp, a terminal

loop and single-stranded RNA segments at both 5' and 3' ends (Ha and Kim, 2014). The pre-miRNA on the other hand is the product of the cleavage of microprocessor complex (Drosha and DGCR8) with ~22 bp at both 5' and 3' ends (**Figure 2.4 (b)**) (Ha and Kim, 2014).

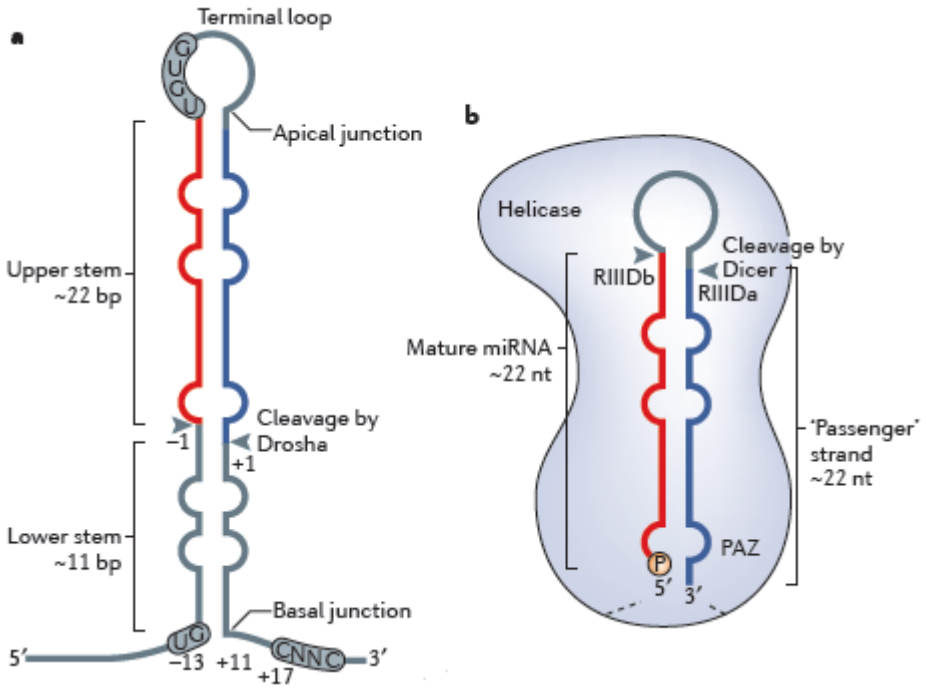


Figure 2.4: (a) The structure of pri-miRNA, consisting of 5', 3' end and terminal loop (stem loop) and the position of cleavage site by Drosha. (b) The structure of pre-miRNA and the cleavage site of Dicer (Ha and Kim, 2014).

Microprocessor complex will recognise and cleaves-off pri-miRNA at the position of approximately 11 bp from the basal junction and 22 bp from the apical junction producing precursor-miRNA (pre-miRNA) leaving two nucleotides overhang in the 3' end (Krol et al., 2010, Ha and Kim, 2014). The pre-miRNA exit the nuclease with the help of exportin 5 (EXP5) which is a Ran-GTP nucleo/cytoplasmic cargo transporter (Ha and Kim, 2014). In the cytoplasm, Dicer cleaves off the stem loop (SL) leaving 3' end (3p) and 5' end (5p) as shown in **Figure 2.4 (b)** (Hutvágner et al., 2001, Ketting et al., 2001). The cleavage formed a small, imperfect dsRNA duplex (miRNA:miRNA*) consisting of mature miRNA strand and 'passenger' strand (miRNA*) (Hutvágner et al., 2001, Ketting et al.,

2001). One of the dsRNA duplex, preferentially with the less stable 5' end is assembled into the RNA-induced silencing complex (RISC) (Schwarz et al., 2003, Khvorova et al., 2003). On the other hand, the 'passenger' strand will be removed immediately after dsRNA duplex is loaded in the pre-RISC complex (Ha and Kim, 2014). The assembling of the less stable miRNA strand portrays the ease of unwinding the dsRNA duplex (He and Hannon, 2004). However, if both of the miRNAs in the dsRNA duplex have similar 5' end stability, it is predicted that both miRNAs have similar frequencies to be established into RISC (He and Hannon, 2004, Schwarz et al., 2003). The miRNA-containing complexes then bind to the 3' UTR of multiple mRNAs as the resident miRNA strand and the target sequence are complementary (Arndt et al., 2009, Iorio and Croce, 2012) causing either translational inhibition or mRNA degradation (Arndt et al., 2009) (**Figure 2.5**).

The choice of targets mRNA being degraded or translationally inhibited is not determined by whether the small silencing RNA originated as an small interfering RNA (siRNA) or miRNA but instead is determined by the degree of sequence complementarity between miRNA and the target mRNA (Billy et al., 2002, Doench et al., 2003). Perfect complementarity results in endonucleocytic cleavage and mRNA degradation while less complementarity leads to repression of protein translation (Reinhart et al., 2000, Hutvágner et al., 2001). Even though translational repression caused by imperfect complementarity seems to be the major mechanism in animals, studies show that several animal mRNAs mediated by mRNA degradation despite being partial complementarity to miRNAs (Wightman et al., 1993, Billy et al., 2002, Lim et al., 2005). The multiple miRNA complementary binding sites have been shown to be the most effective of translational inhibition in metazoan miRNAs (Doench et al., 2003, Lee et al., 1993, Wightman et al., 1993, Reinhart et al., 2000, Abrahante et al., 2003, Lin et al., 2003).

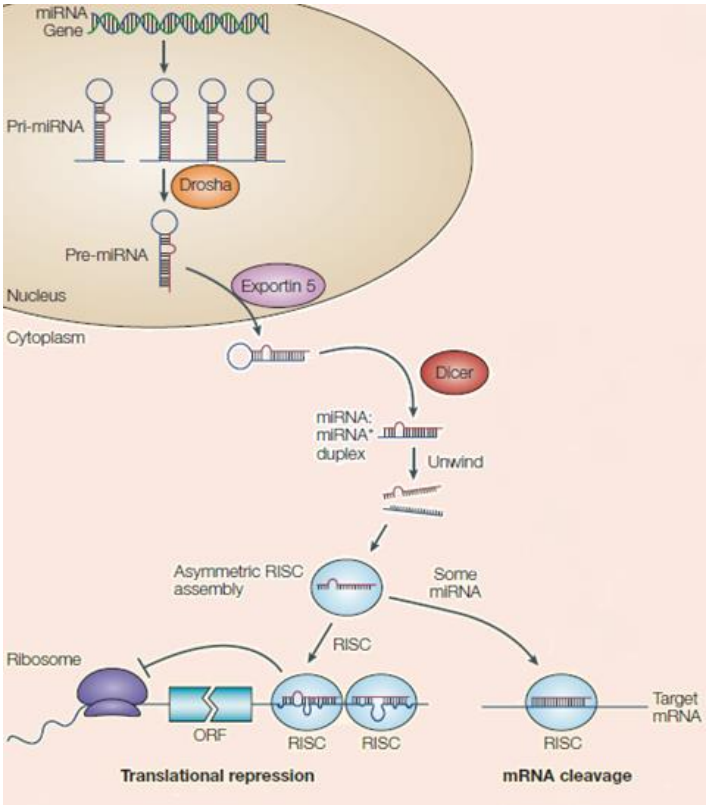


Figure 2.5: miRNA biogenesis and post-transcriptional suppression of miRNA via translational repression or mRNA cleavage (He and Hannon, 2004).

2.53 miRNA130a

The human pre-miR130a (pre-miR130a) with annotation hsa-mir-130a (Accession No. MI0000448, Symbol HGNC:MIR130A) (Geer et al., 2009) is located on the chromosome 11 of *homo sapiens* at 57641198-57641286. The structure of pre-miR130a is made up of 5' (15 - GCU CUU UUC ACA UUG UGC UAC U - 36), 3' (55 - CAG UGC AAU GUU AAA AGG GCA U - 76) and stem loop strand (36 - GUC UGC ACC UGU CAC UAG - 55) (**Figure 2.6**). The miR130a is a vertebrate-specific miRNA and the sequence of mature miR130a-3p are conserved in human, mouse, rat, zebrafish, chicken, clawed frog, opossum, dog, wild boar, chimpanzee, cow, horse and others (Kozomara and Griffiths-Jones, 2014, Kozomara and Griffiths-Jones, 2011, Griffiths-Jones et al., 2006, Griffiths-Jones, 2004). The sequence of miR130a-5p is not conserved and is not detected in all vertebrates but it is expressed in few vertebrates such as human, mouse, rat, chicken and

opossum with slight difference in the sequence (Kozomara and Griffiths-Jones, 2014, Kozomara and Griffiths-Jones, 2011, Griffiths-Jones et al., 2006, Griffiths-Jones, 2004). Studies showed that miRNAs contribute in the regulation of almost all cellular processes and that the changes in their expression observed has been shown to be involved in multi-step carcinogenesis in cancer (Filipowicz et al., 2008). As an evidence, bioinformatics predictions suggested that miRNA may potentially regulate thousands of genes as 30% of all protein coding genes may be affected by miRNAs (Lewis et al., 2005). The miR-130a has been reported to be upregulated in several types of cancer, such as colon cancer, NSCLC, chronic myeloid leukaemia and hepatocellular carcinoma indicating the important roles of miR130a in cancer development, progression and metastasis (Chen et al., 2010, Wang et al., 2010, Suresh et al., 2011, Zhou et al., 2012).

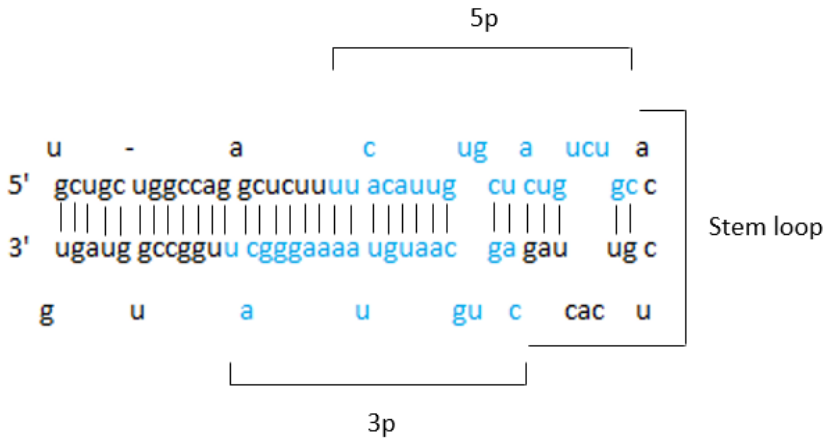


Figure 2.6: Figure shows the structure of pre-miR130a consisting of 5' (5p) end, stem loop and 3' (3p) end (Abdollah et al., 2017).

2.5.4 miR130a expression

The expression of 3p and 5p of miR130a depends on spatial and temporal control of miR130a gene expression in mammalian cells. The expression of miR130a-3p and 5p is selectively high or down regulated in different types of cells and tissues. This has been supported by a study which shows that the selection of mature miRNA strand depends on

cell type, tissue and stimulus (Meijer et al., 2014). The expression of miR130a-3p is down regulated in mixed-lineage leukemic cells (Li et al., 2008), lung squamous carcinoma (Cao et al., 2016), breast cancer (Pan et al., 2015) and colorectal cancer cells (Kara et al., 2015). On the other hand, there were less report on the expression of miR130a-5p especially in cancer studies. A study reported that the expression of miR130a-5p is down regulated in endometrial and endometriotic cells (Braza-Boïls et al., 2015). Another study reported that that both miR-130a-3p and 5p are highly expressed in gestational obesity (Carreras-Badosa et al., 2015).

miRNA has been shown to be involved whether as oncogene miRNA (oncomiR) or tumour suppressor gene in different pathways in cancer. Several studies reported that miR130a is an oncomiR as it was highly expressed in cancer including in gastric cancer cells and cervical cancer cell (Lee et al., 2015, Zhou et al., 2017, Feng et al., 2016, Zhang et al., 2014). In gastric cancer cells, miR130a high expression suppressed the expression of dual-luciferase reporter assays, collapsing response mediator protein 4 (CRMP4) gene and runt-related transcription factor 3 (RUNX3) which enhanced cell proliferation and apoptosis, increases cell colony formation, invasion, migration, and adhesion (Zhou et al., 2017, Lee et al., 2015). In cervical cancer, upregulation of miR130a modulates nuclear factor-*kB* (NF-kB) and tumour necrosis factor-alpha (*TNF-α*) which then target phosphatase and tensin homolog on chromosome 10 (PTEN) and leading to increase in cell proliferation in carcinogenesis of cancer cells (Zhang et al., 2014, Feng et al., 2016).

On the other hand, miR130a acts as tumour suppressor in lung cancer, breast cancer and nasopharyngeal cancer (Lin et al., 2015, Pan et al., 2015, Kong et al., 2018, Chen et al., 2018, Chen et al., 2017). In lung cancer, a study reported that the suppression of miR130a regulates macrophage polarization in NSCLC (Lin et al., 2015). The result showed that mir130a acted as a molecular switch as it expresses at higher level in M1 compared to

M2 macrophages (Lin et al., 2015). Thus, this study concluded that downregulation of miR130a in NSCLC resulted in poor prognosis, increased tumour stage and metastasis (Lin et al., 2015). The expression miR130a was shown to be a tumour suppressor in breast cancer cell line (Pan et al., 2015). As the expression of miR130a was shown to be downregulated, the overexpression of miR130a inhibits cell proliferation, invasion and migration of cancer cells by targeting at ras-related protein Rab-5A (RAB5A) and fos-related antigen 1 (FOSL1) (Pan et al., 2015, Kong et al., 2018). Low expression of miR130a-3p inhibits the expression of transcription regulator gene, BACH2 which suppressing the viability, proliferation, invasion, cell cycle and promotes apoptosis in nasopharyngeal carcinoma cells (Chen et al., 2017).

2.6 Bioinformatics: target gene prediction

In recent years, many bioinformatics platforms are available to analyse large amount of data and to perform comparative analysis of differentially expressed genes. However, there is still no one particular and reliable miRNA target prediction platform that can be used due to computational challenges, scarce miRNA biology knowledge and the inadequate number of experimentally validated target (Betel et al., 2010). As previous study has shown that there are differences in the level of miR130a-3p and 5p expression in different types of cancer cells, thus there is a need to study the role of miR130a-3p and 5p in the regulation of tumorigenesis in cancer. In addition, in this study, the target genes for miR-130a-3p and 5p were identified using combinatorial analysis in miRWalk database. The target genes were then grouped based on function of miR-130a in non-small cell lung cancer (gene ontology and Kyoto Encyclopaedia of Genes and Genomes (KEGG) analysis) and the result were integrated with the mirSVR score for each miR130a-5p and miR130a-3p.

In this study, miRWalk database is used as it is able to combine few databases in order to predict the target genes (Dweep et al., 2014). miRWalk database gathers validated information on miRNA binding sites and also predicts on miRNA and target gene interactions (Dweep et al., 2014). In addition, the validated target gene is updated every month and the predicted target gene is updated for every six months (Dweep et al., 2014). It is important to note that this database shows not only the interaction on the 3' - UTR, but it also includes the other regions of all known genes of the validated target genes (Dweep et al., 2014). Four bioinformatics platforms were used in this study namely miRWalk, the DAVID gene functional classification tool, miRanda-miRSVR analysis tool for gene prediction of miR130a-3p and 5p, and STRING database and protein-protein interaction analysis. DAVID gene functional classification tool is a bioinformatics resource that organised and condensed a list of genes into gene functional groups and classification of genes based on bio-pathways (Huang et al., 2009). miRanda-miRSVR analysis tool on the other hand use the miRanda-predicted miRNA target sites and ranked the them by a down regulation score via mirSVR. This method helps to identify large numbers of experimentally determined non-canonical and non-conserved sites (Betel et al., 2010). Then, the efficiency of miRanda-predicted miRNA target sites will be scored and ranked by miRSVR (Betel et al., 2010). Lastly, STRING database will be used to predict the protein-protein interaction between the predicted target. The STRING database collect and integrate all functional interactions between expressed protein by merging experimented and predicted protein-protein association data in different types of organisms (Szklarczyk et al., 2016). Thus, the bioinformatics analyses were carried out to predict the target genes of miR130a, which will help to understand the function of miR130a and how it interacts with genes identified in associated pathways.