IDENTIFICATION OF NOVEL EPIGENETIC MEDIATORS OF ERLOTINIB RESISTANCE IN NON-SMALL CELL LUNG CANCER

by

Arpita Sushant Pal

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Biological Sciences

West Lafayette, Indiana

May 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Andrea L. Kasinski, Chair

Department of Biological Sciences

Dr. Ourania Andrisani

Department of Basic Medical Sciences

Dr. Elizabeth Tran

Department of Biochemistry

Dr. Harm HogenEsch

Department of Veterinary Administration/
Department of Comparative Pathobiology

Approved by:

Dr. Janice P. Evans

Every challenge in life can be overcome with support and guidance from your loved ones.

For this, I dedicate this work to my parents
Mr. Sushant N. Pal and Mrs. Jolly S. Pal,
my husband,
Dr. Christopher J. Benjamin
and my sisters
Late Ms. Bithika Pal, Ar. Pronita Pal and Ms. Nabamita Pal.

ACKNOWLEDGMENTS

I will take this opportunity to thank all those who have helped me get through this roller-coaster of a journey in graduate school. Firstly, I extend my gratitude to my parents, Mr. Sushant and Mrs. Jolly Pal, for being constant motivators, my greatest mentors and for always having faith in me. Although you are across the globe, there is not one moment I know you are not thinking of me and praying for me. As much I am grateful to my parents, I am thankful to have Christopher Benjamin as my husband. Your unconditional love, support and care, and the little extra push has got me to the end of this journey. My sisters, Pronita and Nabamita Pal, without you two, I could not have made it through all these years. I want to thank you for not just hearing me out in my times of distress, but for sharing the most joyful moments of my life, despite the distance between us. I especially want to acknowledge my beloved oldest sister, late Ms. Bithika Pal and my favorite aunt, late Mrs. Radha Pal, thank you both for watching over me from the heavens above. And a big thank you to my uncle, Mr. Samaresh Pal, and my grand-parents Mr. Sudhir Chandra and Mrs. Biva Ghosh, late Mr. Nityagopal and late Mrs. Tara Rani Pal for your blessings and constant support. I hope I can make you all proud by the little accomplishments that I achieve.

I am sincerely appreciative of the Benjamins for welcoming me into their family, and becoming my new family in the U.S. Thank you Ms. Aleta Benjamin, Mr. Mark and Mrs. Teressa Benjamin and Mr. Daryl Freeman for keeping me in your prayers, and checking up on me every so often.

In the professional front, I am truly thankful to Dr. Andrea Kasinski for giving me the opportunity to enter her lab as a graduate student. My research started off with conducting two screens, not knowing what the end result would be, but I thank you for trusting in me and entering into a study that was away from the lab's expertise. Under your advice and direction, I have learnt to approach my work with a sense of ownership. Your critique in both research and writing, and your guidance in teaching and presenting my work have helped me develop into the researcher and communicator that I am today. I admire your zeal and strength in pursuit of science, and I hope that my approach to science upon leaving this lab is as passionate and ingenious as yours.

I also want to thank my thesis committee members who have been a reliable source of instruction all along. Dr. Ourania Andrisani, your mentorship has been invaluable to me since I was accepted

into Purdue in 2013. Thank you for your advice, encouragement, and moral support throughout my graduate career. Dr. Elizabeth Tran and Dr. Harm HogenEsch, I am grateful to you for your critical assessments during each committee meeting. I deeply appreciate the guidance from such a diverse advisory committee that has trained me to think critically and confidently navigate the project into an uncharted direction.

Without the help and support that I have received from the members of Kasisnki Lab, this journey would only have been tougher. I want to extend a special thanks to Alejandra Agredo for getting on board with me on the major project of my thesis. With her incessant help in conducting experiments since June of 2018, the project has gathered pace and many exciting results have come into fruition. Among a few past and the other present members, I want to extend my thanks to Esteban Orellana, Humna Hassan, Ikjot Singh Sohal, Manvir Bains and the rest of the team for making the lab a conducive and fun environment to work in.

Purdue has provided an exceptional platform to conduct experiments with instruments, technology, and facilities available to very few, but has also given me the opportunity to interact with some outside of my research that I can call my friends for life. I want to thank Supriya Dharkar, Benjamin Seelbinder, Eshaan Mathew, Ankita Thawani, Adhiraj Mathur, Plabita Taludkar and a few more for making my experience at Purdue a memorable one. Among non-Purdue affiliated folks, I am thankful to Renita Jane Pinto for being by my side through thick and thin, almost forever now.

Lastly, I want to acknowledge the newest and upcoming member of the Pal-Benjamin family who has been very supportive since I started writing my thesis, and a true inspiration to finishing graduate studies and make more dreams come true.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGO	Argonaute
ANOVA	Analysis of variance
ATCC	American Type Culture Collection
ATP	Adenosine-Triphosphate
BCAA	branched-chain amino acid
BCKDH	branched chain α-keto acid dehydrogenase complex
BCL2	BCL2 Apoptosis Regulator
C/EBP	CCAAT Enhancer Binding Protein
CAGRs	Cancer-Associated Genomic Regions
CARIP-seq	chromatin-associated RNA immunoprecipitation and sequencing
CASP8	Caspase 8
CCND1	Cyclin D1
CDS	coding sequence
ChIP	chromatin will be immunoprecipitated
ChIP-qPCR	chromatin will be immunoprecipitated-quantative PCR
CLL	Chronic Lymphocytic Leukemia
CMV	Cytomegalovirus
c-Myc / MYC	MYC Proto-Oncogene, BHLH Transcription Factor
COMET	Correlated-to-MET
COPA	Cancer Outlier Profile Analysis
Cre-loxP	Cre recombinase and a loxP recognition site
CRISPR	clustered regularly interspaced short palindromic repeats
cRNA	CRISPR RNA
CT	computerized tomography
CtRP	Cancer Therapeutics Response Portal
DDR	DNA damage repair
DGCR8	DGCR8 Microprocessor Complex Subunit
DICER	Dicer 1, Ribonuclease III
DMSO	dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DNMT1	DNA Methyltransferase 1
DNMT3	DNA Methyltransferase 3
DOX	doxycycline
DROSHA	Drosha Ribonuclease III
DTP	Developmental Therapeutics Program
ECas9	EKVX cells stably expressing Cas9
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal Growth Factor

EGFL7	EGF Like Domain Multiple 7
eGFP	enhanced Green Fluorescent Protein
EGFR	Epidermal Growth Factor Receptor
EKVX-pmiR	pmiRGLO expressing EKVX cells
EMT	epithelial to mesenchymal transition
ER	Estrogen Receptor
FBS	Fetal Bovine Serum
Flp/FRT	Flippase/Flp recombinase target
FOXA1	Forkhead Box A1
G418	Geneticin or Neomycin
GAL4	galactose-responsive transcription factor
GAL4-UAS	GAL4/upstream activating sequence (UAS) system
GeCKO	Genome-scale CRISPR Knock-Out
GEPIA	Gene Expression Profiling Interactive Analysis
GI50	50% Growth Inhibitory Concentration
GI75	75% Growth Inhibitory Concentration
GOI	Gene of Interest
GSTP1	glutathione S-transferase P1
GTEx	Genotype-Tissue Expression
H322M-pmiR	pmiRGLO expressing H322M cells
H4K20	Histone H4- Lysine 20 H4K20
HBEC	Human Bronchial Epithelial Cells
HCl	hydrochloric acid
HDR	homology-directed repair
HGF	hepatocyte growth factor
HGFR	hepatocyte growth factor receptor
HMGA2	High Mobility Group AT-Hook 2
ICW	In-cell western
IF	Immunofluorescence
IgG	Immunoglobulin G
IP	immunoprecipitation
IPA	Ingenuity Pathway Analysis
KMT2D/ MLL2	histone-lysine N-methyltransferase 2D
KMT5C	Lysine Methyltransferase 5C
KRAS	Kirsten rat sarcoma 2 viral oncogene homolog
LCN2	Lipocalin-2
LDCT	low-dose computed tomography
lncRNA	long non-coding RNA
LUAD	lung adenocarcinoma
LUSC	lung squamous cell carcinoma

MAGeCK-VISPR	Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout-VISualization of crisPR screens
mdig	Mineral Dust Induced Gene Protein/ Myc-Induced Nuclear Antigen
MET	mesenchymal-epithelial transition
MET (the protein)	hepatocyte growth factor receptor
miRISC	miRNA-induced silencing complex
miRNAs	microRNAs
MOI	multiplicity of infection
MRI	magnetic resonance imaging
mTOR	mammalian target of rapamycin
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCI	National Cancer Institute
Neo	Neomycin
NF1	Neurofibromin 1
NHEJ	non-homologous end-joining
NP-40	Nonidet P-40
NSCLC	non-small cell lung cancer
OE	Overexpression
OncomiR	Oncogenic miRNA
OPA3	Outer Mitochondrial Membrane Lipid Metabolism Regulator
ORF	Open Reading Frame
OSCC	oral squamous cell carcinoma
PAM	protospacer adjacent motif
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDCD4	programmed cell death 4
PET	positron emission tomography
PI3K/AKT	phosphatidylinositol 3-kinase/Protein kinase B
PIPES	piperazine-N,N′-bis(2-ethanesulfonic acid)
pre-miRNA	precursor miRNA
pri-miRA	primary miRNA
PTEN	phosphatase and tensin protein
PTENP1	PTEN homologous pseudogene
PVDF	polyvinylidene difluoride
qRT-PCR	Quantitative real time PCR
RAS/MAPK	RAS/mitogen-activated protein kinase
RIPA buffer	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
ROS	reactive oxygen species
RPMI medium	Roswell Park Memorial Institute Medium

rRNA	ribosomal RNAs
SCC	squamous cell carcinoma
SCLC	small cell lung cancer
SDS	Sodium deoxycholate
sgRNA	small guide RNA
SH2	Src-homology 2 domain
sicont	Silencer Select Negative Control
siLUC2	siRNA targeting Firefly Luciferase
siRNA	silencing RNA
snoRNA	small nucleolar RNAs
snRNA	small nuclear RNAs
SRB assay	Sulforhodamine B colorimetric assay
STAT3/5	signal transducer and activator of transcription 3/5
SUV420H2	Su(Var)4-20 Homolog 2
TANRIC	The Atlas of Noncoding RNAs in Cancer
TCA	tricarboxylic acid
TCGA	The Cancer Genome Atlas
Tet	Tetracycline
TKI	EGFR tyrosine kinase inhibitor
TP53	tumor protein p53
TPM	Transcripts per million
tracrRNA	trans-activating crRNAs
TRBP	Transactivation-Responsive RNA-binding protein
tRNA	transfer RNA
tz	time zero
UT	Untransfected or untreated
UTR	untranslated region
γ radiation	gamma radiation
γ-H2AX	Gamma H2A.X Variant Histone

ABSTRACT

Lung cancer is the third most prevalent cancer in the world; however it is the leading cause of cancer related deaths worldwide. Non-small cell lung cancer (NSCLC) accounts for ~85% of the lung cancer cases. The current strategies to treat NSCLC patients with frequent causal genetic mutations is through targeted therapeutics. Approximately 10-35% of NSCLC patient tumors have activated mutations in the Epidermal Growth Factor Receptor (EGFR) resulting in uncontrolled cellular proliferation. The standard-of care for such patients is EGFR-Tyrosine Kinase Inhibitors (EGFR-TKIs), a class of targeted therapeutics that specifically inhibit EGFR activity. One such EGFR-TKI used in this study is erlotinib. Following erlotinib treatment, tumors rapidly regress at first; however, over 50% of patients develop erlotinib resistance within a year post treatment. Development of resistance remains to be the major challenge in treatment of NSCLC using EGFR-TKIs such as erlotinib.

In approximately 60% of cases, acquired erlotinib resistance in patients is attributed to a secondary mutation in EGFR, whereas in about 20% of cases, activation of alternative signaling pathways is the reported mechanism. For the remaining 15-20% of cases the mechanism of resistance remains unknown. Therefore, it can be speculated that the common methods used to identify genetic mutations in tumors post erlotinib treatment, such as histologic analysis and genetic screening may fail to identify alterations in epigenetic mediators of erlotinib resistance, also including microRNAs (miRNAs). MiRNAs are short non-coding RNAs that post-transcriptionally negatively regulate their target transcripts. Hence, in this study two comprehensive screens were simultaneously conducted in erlotinib sensitive cells: 1) a genome-wide knock-out screen, conducted with the hypothesis that loss of function of certain genes drive erlotinib resistance, 2) a miRNA overexpression screen, conducted with the hypothesis that certain miRNAs drive the development of erlotinib resistance when overexpressed. The overreaching goal of the study was to identify novel drivers of erlotinib resistance such as microRNAs or other epigenetic factors in NSCLC.

The findings of this study led to the identification of a tumor suppressive protein and an epigenetic regulator, SUV420H2 (KMT5C) that has never been reported to be involved in erlotinib resistance.

On the other hand, the miRNA overexpression screen identified five miRNAs that contribute to erlotinib resistance that were extensively analyzed using multiple bioinformatic tools. It was predicted that the miRNAs mediate erlotinib resistance via multiple pathways, owing to the ability of each miRNA to target multiple transcripts via partial complementarity. Importantly, a correlation between the two screens was identified clearly supporting the use of two simultaneous screens as a reliable technique to determine highly significant miRNA-target interactions. Overall, the findings from this study suggest that epigenetic factors, such as histone modifiers and miRNAs function as critical mediators of erlotinib resistance, possibly belonging to the 15-20% of NSCLC cases with unidentified mechanisms involved in erlotinib resistance.

CHAPTER 1. INTRODUCTION

1.1 Cancer:

Cancer is a disease caused by abnormal changes in a cell that provides the cell with a growth and survival advantage, and the ability to disseminate to different organs. However, the development of cancer is described as a progressive multi-step process. First, a normal cell incurs genetic changes that allows the cell to evade cell death, growing and transforming into a non-invasive mass. The mass further undergoes uncontrollable proliferation, invading into the surrounding tissues and lymph nodes, followed by traversing through the vasculature to metastasize to distant organs. The genetic and physiological changes that a normal cell undergoes to transform into a cancer cell are described as hallmarks of cancer (1).

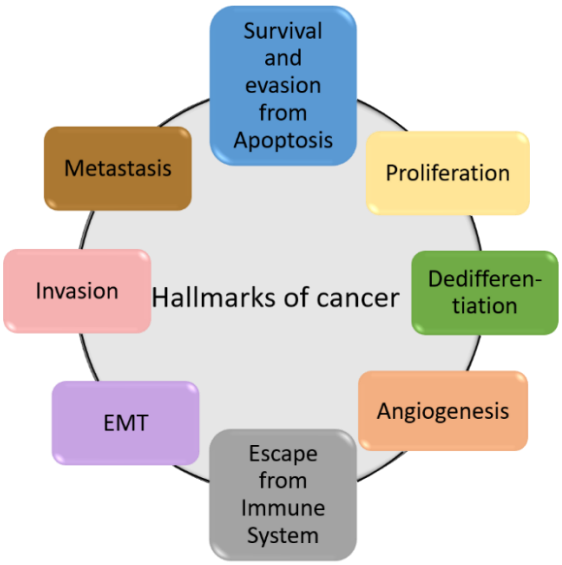


Figure 1.1. Hallmarks of cancer, depicting the various changes involved in transforming a normal cell into a cancerous state (Adapted from (1)).

In the process of cancer development, as described above, acquiring genetic changes in a normal cell initiates the onset of hallmarks of cancer. Genomic instabilities and chromosomal aberrations are common triggers initiating wide-ranging changes to the genome (2–5). Additionally, prevalent genetic alterations of particular importance occur in proto-oncogenes and tumor-suppressive genes. Proto-oncogenes are genes that function to maintain normal cellular processes such as survival or proliferation of a cell, until triggered by activating mutations or amplification, resulting in the

generation of an oncogene. Whereas in a normal cell, a tumor-suppressive gene functions to protect the cell against development of tumors. However, during the onset of cancers, tumor suppressive genes are mutated that result in loss of function of their tumor-suppressive activity. Multiple oncogenes and tumor-suppressive genes have been identified that function as drivers of cancer. A well-studied tumor suppressor gene tumor protein p53 (*TP53)*, which codes for the protein p53 is intrinsically activated following DNA damage resulting in cell cycle arrest and apoptosis, essentially serving as the guardian of the genome. *TP53* is therefore an essential gene required to maintain genome integrity, and perhaps not surprisingly, is the most frequently altered gene in human cancers (6–9). Mutations in *TP53* result in loss of function of wild-type (WT) p53 thereby initiating the onset of cancers. Most mutations in p53 abrogate its wild type activity while simultaneously gaining additional functions that result in the formation of more aggressive tumors (9). The most common hot-spot mutations in p53 include missense mutations at R249S, R175H, R248W, and R273H (9,10). Additionally, a common driver mutation in a proto-oncogene that results in the generation of an oncogenic protein occurs in the *KRAS* (Kirsten rat sarcoma 2 viral oncogene homolog) gene. Mutant KRAS drives the development of multiple cancers (7,11). The normal cellular function of KRAS is to respond to external stimuli of cell growth, proliferation and chemotaxis via activation of multiple downstream signaling pathways (12,13). Whereas in cancers, mutant KRAS constitutively signals downstream, independent of upstream activators resulting in uncontrolled growth and proliferation (12,14). However, co-mutations of KRAS with another gene, such as TP53, STK11 or KEAP1 further potentiates KRAS-driven tumorigenicity (13,15). Apart from KRAS and TP53 mutations, mutations in additional drivers of cancer have been identified by studies conducted using the enormous publicly available cancer datasets contained in The Cancer Genome Atlas (TCGA) (6,7,11,16).

In addition to acquiring genetic alterations, cancer cells are also capable of modulating their surroundings, known as the microenvironment. Changes in the microenvironment can favor growth, proliferation, and eventually metastasis of the cancer (17,18). Endogenous genetic alterations in combination with microenvironment modulation accounts for the formation of aggressive tumors (19,20). Such adaptations and crosstalk between cancer cells and cells of the microenvironment makes cancer evermore complex, and difficult to treat. However, current research has advanced our understanding of a few distinct characteristics of various cancers, including specific genetic backgrounds and mediators of tumors in the microenvironment (7,21–

23). This knowledge has facilitated the development of several treatment strategies, but one drawback that results in poor clinical outcomes, post-treatment of cancer, is that most patients develop resistance to therapy and suffer relapse. This is mainly attributed to the manifestation of a cancer hallmark wherein cancer cells gain the ability to evade apoptosis, conferring them the ability to survive in the presence of the therapeutic agent (24–26). Despite the advancements made in understanding the molecular basis of cancer, it remains to be a major health concern worldwide and is the second most common cause of mortality in the United States. Hence, it is conceivable to claim that cancer is still an elusive disease with largely ineffective treatment strategies. Therefore, to improve clinical outcomes, current research has galvanized towards developing effective treatment strategies, and unraveling the oncogenic mechanisms to prevent the development of resistance to the prevailing therapeutics.

1.2 Lung cancer:

Cells of the lungs that deviate from normal growth and proliferation due to spontaneous genetic changes, or alterations due to external stimuli such as smoking, undergo transformations that enable these cells to transform into pre-cancerous cells. Upon accruing more abnormal changes described as the hallmarks of cancer (**Figure 1.1**), these cells are converted into a neoplastic mass, ultimately referred to as lung cancer. Lung tumors may interfere with the normal function of the lungs, which is to provide oxygenated blood to the entire body. Moreover, at later stages, lung tumors can spread and disseminate to parts of the body such as the brain, bone, and liver impairing additional organ and bodily functions.

1.2.1 Lung cancer etiology:

Lung cancer develops when normal lung cells become mutated due to either chronic exposure to cancer causing substances, called carcinogens, or due to genetic predisposition that lead to altered expression of certain genes. Exposure to carcinogens and spontaneously altered genes, along with other factors that predispose a person to developing lung cancer are classified as risk factors. The most common carcinogens involved in promoting lung cancer include tobacco smoke, radon, asbestos and other such harmful chemicals. Other common risk factors include unhealthy diet, air pollution, predisposition due to age, family history or previous lung conditions (27,28). Therefore,

in order to prevent the development of lung cancer, it is necessary to avoid the preventable risk factors and take preventative measures such as screening for an early diagnosis of cancer.

1.2.2 Lung cancer diagnosis and staging:

Lung cancer is the second most commonly diagnosed cancer in both men and women, after prostate cancer in men and breast cancer in women (29). However, it is the leading cause of cancer mortality worldwide. Lung cancer mortality is attributed to the diagnosis of cancer at late stages. Only at late stages, lung cancer manifests symptoms such as coughing of blood, weight loss, loss of appetite, chest pain and shortness of breath. Patients displaying such symptoms are further examined and tested for lung cancer (28). Currently, chest X-rays, computerized tomography (CT) scan, magnetic resonance imaging (MRI) scan, or positron emission tomography (PET) scan is prescribed to detect tumors. Once the tumor/s are located, biopsy and biomarker testing are conducted to evaluate the stage of cancer (**Table 1**). More recently, genetic testing is being conducted to identify casual mutations in lung cancer cells, and subsequently determine appropriate treatment regimens (27,28).

Although late-stage diagnosis of lung cancer is common, an early-stage diagnoses is beneficial for the overall survival of patients. In some rare cases, patients are diagnosed with early-stage lung cancer, while under treatment for unrelated or non-cancerous lung diseases (27). It has become common practice for high risk individuals (due to familial history) to screen for lung cancer markers in order to diagnose lung cancer before symptoms manifest. Screening technologies such as low-dose computed tomography (LDCT) scanning of the chest, or sputum cytology are now routinely utilized (28,30). Lung cancer patients that are diagnosed at early stages (**Table 1**), i.e. when the tumor is present only at the primary site (localized), have a favorable 5-year survival rate relative to patients diagnosed at late stages, i.e. when the cancer has spread to lymph nodes (regional), or metastasized (distant) (**Figure 1.2**) (31). Therefore, there is a need to diagnose and identify lung cancer at early stages or develop better treatment strategies to benefit patients diagnosed with lung cancer at late stages.

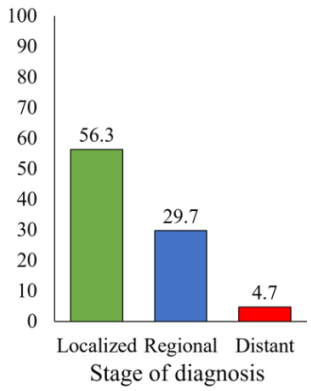


Figure 1.2. Five-year survival rate of lung cancer is dependent on stage of diagnosis of cancer. Patients with localized stage of cancer display a better 5-year overall survival relative to those with late stages, regional and distant cancers (compiled from (31)).

1.2.3 Histological origin of lung cancer:

Different cell types compose various parts of the lung. The pharyngeal airway leads into the trachea that is composed of basal, goblet (mucosal), clara, and ciliated epithelial cells. Downward, as the trachea divides into bronchi, neuroendocrine cells cluster in the lining of the epithelium. The bronchi form bronchioles, primarily composed of clara and ciliated epithelial cells, which further differentiate into alveolar sacs. Alveolar sacs are composed of specialized thin epithelial cells called pneumocytes, responsible for gas exchange (32). Based on the cell of origin and histology of the tumor lung cancer has been divided into two major subtypes: 1) small cell lung cancer (SCLC) that originate from neuroendocrine cells in bronchi, and 2) non-small cell lung cancer (NSCLC) which is composed of a variety of cells and originates from various locations in the lung (33,34), described in **Figure 1.3**. SCLCs are aggressive cancers that metastasize to distant organs such as the liver, bone, and brain (35), but are predominantly diagnosed in smokers. Whereas NSCLC comprises the majority of lung cancer cases – ~ 85% of diagnosed lung cancer patients. Thus NSCLC cases are further sub-classified into 40% adenocarcinoma, 25% to 30% squamous cell carcinoma (SCC), and 10% to 15% large cell carcinomas (**Figure 1.4**)(31,36).

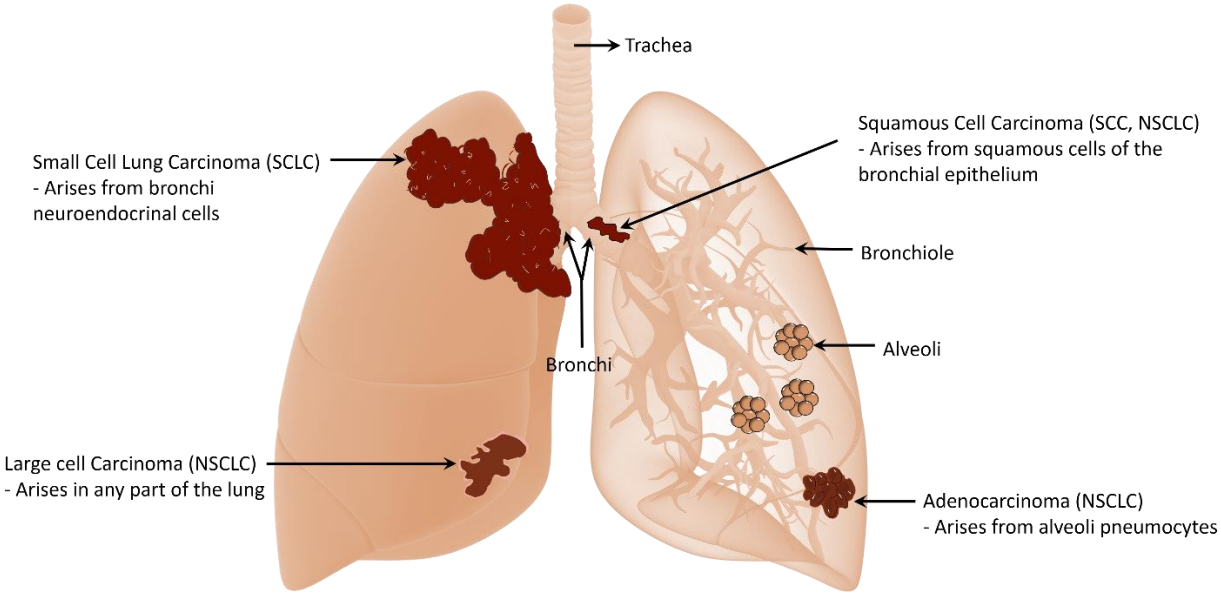


Figure 1.3. Histological origin of lung cancer. Small cell lung cancer (SCLC) originates from neuroendocrinal cells in bronchi whereas non-small cell lung cancer (NSCLC) originates from a variety of cells in the lung. Squamous cell carcinoma (SCC) arises from squamous cells of the bronchial epithelium (33,34,37). The origin of adenocarcinoma is from alveoli pneumocytes. Large cell carcinoma originates from cell types other than cells giving rise to SCC or adenocarcinoma (33,34). Illustration created using graphics from Library of Science & Medical Illustrations (http://www.somersault1824.com/science-illustrations/)

Table 1.1. Stages of lung cancer describing tumor characteristics and extent of spread of tumor (compiled from (27)).

Stage		**Cell/Tumor characteristics**	**Extent of spread**
Stage 0		Malignant cells, carcinoma *in situ*	Localized
Stage 1	**1A**	Tumor is up to 3 cm	Spread into the lung inner lining, but not to the lymph nodes
	1B	Tumor is 3 - 4cm	
Stage 2	**2A**	Tumor is 4 - 5cm	Spread into the lung inner lining, but not to the lymph nodes
	2B	Tumor is 5 - 7cm	Spread into the lung inner lining and nearby sites, but not to the lymph nodes, secondary tumors in the same lobe
		Tumor is 5cm	Spread into the lung inner lining and nearby sites, but not to the lymph nodes, lung has collapsed or is swollen
Stage 3	**3A**	Tumor is up to 7 cm	Locally advanced, but did not spread to the lymph nodes, secondary tumors appear in the same lung but different lobe
		Tumor is less than 5 cm	Locally and advanced to the lymph nodes, secondary tumors appear in the same lung but different lobe
	3B	Tumor is more than 5 cm	Locally and advanced to the lymph nodes, secondary tumors appear in the same lung but different lobe
		Tumor is less than 5 cm	Advanced to the lymph nodes near other lung, secondary tumors may appear in the main bronchus, lung may have collapsed or is swollen
		Tumor is more than 5 cm	
Stage 4	**4A**	Any size tumor	Tumor has not spread to lymph nodes, but metastasized to other lung and one site outside chest the area
	4B	Any size tumor	Tumor may or may not have spread to lymph nodes, but metastasized to multiple sites outside chest the area

1.2.4 Pathology of lung cancer

The development of lung cancer occurs when a normal lung cell suffers a series of progressive changes, known as preneoplastic or precancerous changes. The preneoplastic changes in the lung epithelium include hyperplasia, squamous metaplasia, squamous dysplasia, and carcinoma *in-situ* (CIS) leading to the formation of precancerous lesions (38,39). Precancerous lesions appear on the top layer of the bronchi, bronchioles, or alveoli, without invading deeper. However, when the cells invade into deeper layers of the site of origin, lung cancer develops and the cells acquire an ability to metastasize (39). SCLC develops as the neuroendocrinal cells of the bronchial epithelium undergo hyperplasia, bypassing the formation of precancerous lesions. However, NSCLC develops progressively via sequential preneoplastic changes (38,39). For example, SCC develops as the normal lung epithelium undergoes hyperplasia followed by metaplasia or dysplasia that results in CIS, which acts as a precursor of SCC. Lung adenocarcinoma initiates as the alveolar epithelium become hyperplastic, regarded as atypical adenomatous hyperplasia, progressing into the formation of a non-invasive CIS, i.e. bronchioalveolar carcinoma that eventually develops into invasive lung adenocarcinoma (38).

1.3 Non-small cell lung cancer (NSCLC):

NSCLC is a type of lung cancer composed of the various aforementioned sub-groups of cancers including adenocarcinomas, squamous cell carcinomas, and large cell carcinomas (**Figure 1.3, 1.4**). NSCLC classification is typically based on histology and anatomical location, but in light of recent understanding of genetic drivers, a more appropriate classification system has been developed, described in the following section.

1.3.1 NSCLC classification based on genomic profiling:

Although is still accepted, the traditional classification of NSCLC based on histological origin is now considered a simplified classification that does not appropriately address heterogeneity of tumors. Therefore, with the advent of genomic profiling of NSCLC tumors, genetic changes such as gene amplifications or loss of function of genes, or chromosomal rearrangements have been identified and are used for more precise classification (**Figure 1.5**). This has led to the

identification of genes with certain mutations that are referred to as “driver mutations” that are causal of NSCLC, and are therefore the basis of NSCLC classification (27,31,40).

1.3.1.1 Driver mutations in proto-oncogenes:

Driver mutations or genetic changes in NSCLC are mainly contained in proto-oncogenes, which are dormant oncogenes performing normal cellular functions, such as KRAS, EGFR, MEK, MET, HER2 (**Figure 1.5**) (11,41–45). When proto-oncogenes become activated, they function as oncogenes – genes that promote cancerous phenotypes. The pathways activated by an oncogene allows the cell to attain survival and uncontrollable growth abilities, evading apoptotic and necrotic signals.

One example of a protooncogene activated in ~ 2% of NSCLC is MET (mesenchymal-epithelial transition), also known as hepatocyte growth factor receptor (HGFR) (**Figure 1.5**) (46–48). MET is a receptor tyrosine kinase, expressed on the surface of epithelial cells that activates several signaling pathways leading to growth, proliferation, motility, and epithelial to mesenchymal transition (EMT) upon binding to its ligand, hepatocyte growth factor (HGF). The normal function of MET is to activate downstream effector pathways such as phosphatidylinositol 3-kinase/AKT (PI3K/AKT), mammalian target of rapamycin (mTOR), signal transducer and activator of transcription 3/5 (STAT3/5), and RAS/mitogen-activated protein kinase (RAS/MAPK). However, in NSCLC, MET overexpression occurs due to a mutation in its juxta-membrane domain that results in exon 14 skipping (METex14). METex14 prevents MET degradation, hence overabundance leading to constitutive activation of its downstream signaling pathways (49–51).

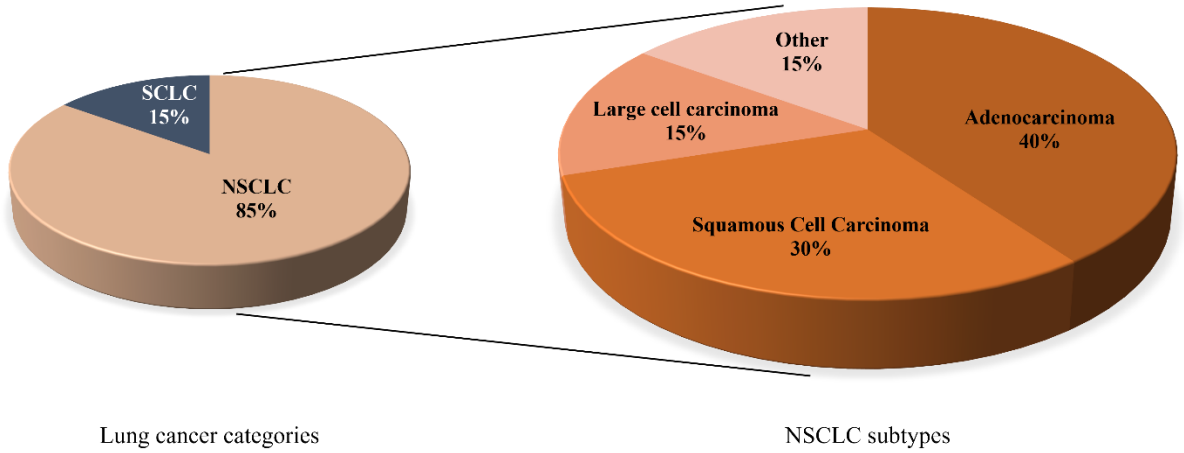


Figure 1.4. Histological classification of lung cancer. Lung cancer is classified into two major categories, of which 15% of cases are small cell lung cancer (SCLC) and 85% are non- small cell lung cancer (NSCLC), represented in a pie-chart (left). NSCLC is further categorized into three main subtypes (right). Adenocarcinoma comprises of 40%, squamous cell carcinoma (SCC) approximately 25% to 30% of NSCLC cases, and large cell carcinomas is approximately 10% to 15% of all NSCLC cases. The remaining 15% is categorized as "other" or "unspecified due to mixed cellular characteristics" (compiled from (31,34)).

1.3.1.2 Loss of tumor-suppressor genes:

Tumor suppressor genes act as gatekeepers to prevent tumor formation, many of which normally function to prevent cell division in the presence of genomic mutation or to induce cell death upon receiving appropriate signals. One such tumor suppressor that is frequently lost in NSCLC is PTEN, which encodes for phosphatase and tensin protein. Normal function of PTEN is to prevent cells from rapidly growing and proliferating through inhibiting the PI3K/Akt pathway, and in turn triggering apoptosis (52,53). Therefore, loss of PTEN inadvertently leads to survival and growth via augmented PI3K/Akt signaling in NSCLC. However, loss of PTEN in NSCLC is not typically associated with genomic inactivating mutations. Instead, reduced PTEN expression, observed in 4-8% of NSCLC patients is due to epigenetically regulation (**Figure 1.5**). The promoter of PTEN and the PTEN homologous pseudogene, *PTENP1* both enhance PTEN activity. However, in NSCLC *PTEN* and *PTENP1* are regulated via hypermethylation, leading to loss of PTEN function thereby driving NSCLC via uncontrolled growth and evasion of apoptosis (24,52,54). Additional aspects of epigenetic changes and the role of epigenetics in mediating cancer is further described in Section 1.6.

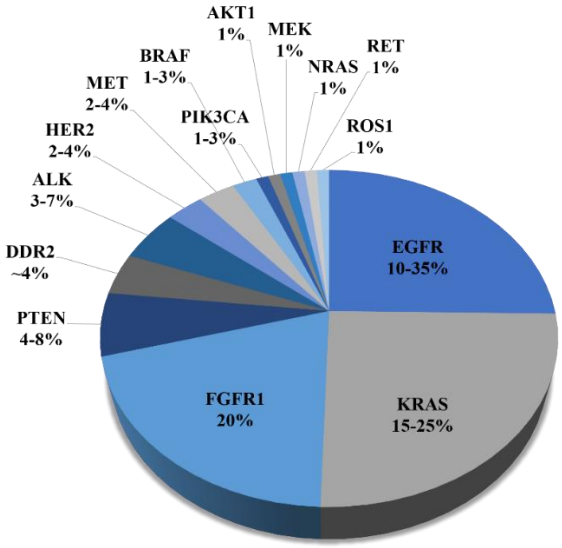


Figure 1.5. Profile of mutations and genetic alterations in NSCLC. Genetic alterations identified in NSCLC tumors are depicted in a pie-chart. The most commonly mutated genes in NSCLC, such EGFR, KRAS are also known as drivers of NSCLC (compiled from (27,31,40)).

1.3.2 Treatment of NSCLC:

Following lung cancer diagnosis, patients are treated with one or more of the current treatment interventions, depending on 1) stage of the disease, and 2) overall health of the patient. These interventions include surgery, chemotherapy, radiotherapy, targeted therapy, and immunotherapy. If the tumor is diagnosed in the early stages, such that it is localized at the primary site, the tumor is resected via surgery, usually followed by radiotherapy and/or conventional chemotherapies. However, in the case of most NSCLC patients with cancer at advanced stages, the tumor is profiled to identify a driver mutation followed by treatment with targeted therapy (27,28). But currently the paradigm is shifting, and more patients are being administered a targeted agent regardless of profiling, rather than the globally acting chemotherapies, which are discussed in detail in the following sections.

1.3.2.1 Chemotherapy versus targeted therapies:

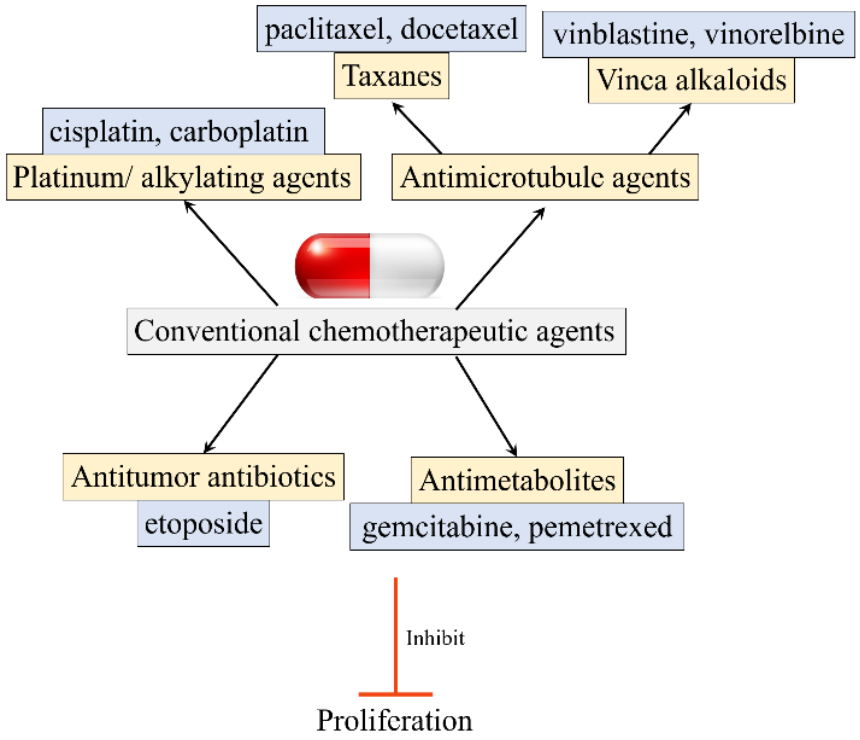
Conventional chemotherapies used to treat NSCLC patients are cytotoxic agents that target rapidly diving cells, which is a characteristic of cancerous cells. These agents interact with the DNA of proliferating cells, inhibiting DNA synthesis and/or repair mechanisms. For patients whose tumors are localized and are thus candidates for surgery, post-surgical resection is followed with either orally or intravenously administered first-line chemotherapy treatments (i.e. initial treatment). Chemotherapy combined with surgical resection of tumors is also known as adjuvant therapy. When the tumor is localized, chemotherapy is used as a precautionary mechanism to prevent cancer relapse. However, in cases when the tumor is too large to be surgically resected without additional interventions, chemotherapy is administered prior to surgery, also known as neo-adjuvant therapy (55,56). Chemotherapeutic agents commonly used in NSCLC are categorized into four types based on their mechanism of action: 1) cisplatin and carboplatin (platinum/alkylating agents), 2) gemcitabine and pemetrexed (antimetabolites) 3) etoposide (antitumor antibiotics), 4) paclitaxel and docetaxel (taxanes/ antimicrotubule agents), vinblastine and vinorelbine (vinca alkaloids/ antimicrotubule agents) (**Figure 1.6**).

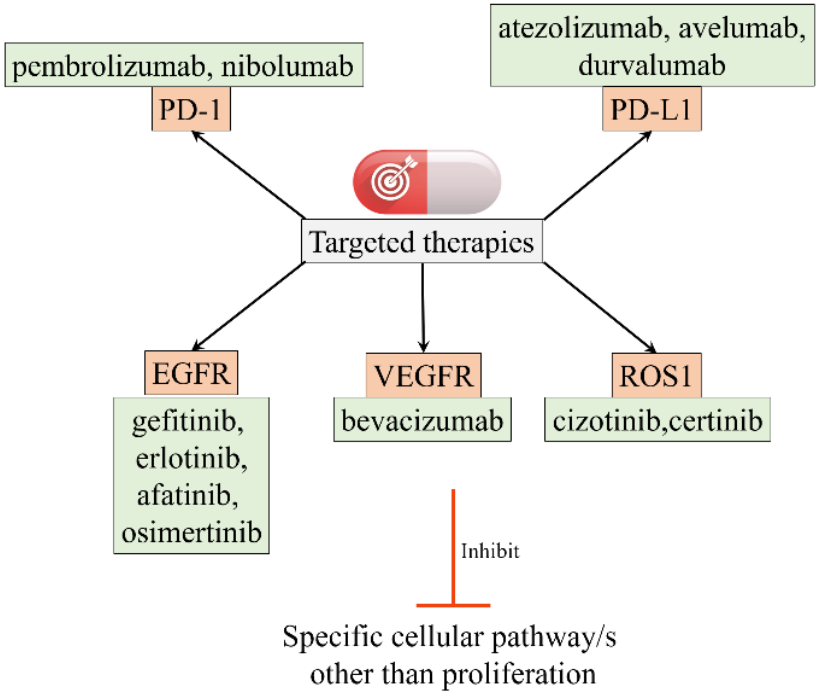
In cases of relapse post-chemotherapy, patients are administered with the standard-of-care as a second-line of treatment (27,57). Currently, the standard-of-care for NSCLC patients is a targeted agent, especially if the tumor harbors a driver mutation/alteration (**Figure 1.5**). Targeted therapies specifically inhibit activity of the gene that is driving cell survival and proliferation, thereby circumventing the toxicities associated with chemotherapies that act as global DNA damaging agents in all proliferating cells (**Figure 1.6**) (27,55–57). One of the fist targeted therapies generated was against an activating mutation in the Epidermal Growth Factor Receptor (EGFR), a receptor tyrosine kinase that is mutated in ~10-35% of NSCLC patients (**Figure 1.5**). EGFR tyrosine kinase inhibitor (TKI) treatment resulted in an improvement in the overall and progression free survival of a specific subset of patients (58,59). This observation led to an exponential increase in research and establishment of clinical trials to develop and test multiple EGFR-TKIs that specifically target tumors with constitutively active EGFR (58,59).

1.4 Epidermal Growth Factor Receptor (EGFR):

Epidermal Growth Factor Receptor (EGFR) is a driver gene of NSCLC, accounting for 15-35% of mutations or alterations in NSCLC cases (**Figure 1.5**). Due to the importance of EGFR in the body of work presented in this thesis, the normal function of EGFR in cells followed by its role in cancer, and effect of TKIs on EGFR is explained in detail in the following sections.

Figure 1.6. Chemotherapeutic agents and targeted therapies used in NSCLC treatment. Classification of various chemotherapeutic agents based on cytotoxic characteristics, resulting in inhibition of proliferation (left). Targeted therapies targeting drivers of NSCLC resulting in inhibition of specific pathways apart from proliferation (right) (adapted from (57,60)).


paclitaxel, docetaxel
Taxanes
vinblastine, vinorelbine
Vinca alkaloids
cisplatin, carboplatin
Platinum/ alkylating agents
Antimicrotubule agents
Conventional chemotherapeutic agents
Antitumor antibiotics
etoposide
Antimetabolites
gemcitabine, pemetrexed
Inhibit
Proliferation


pembrolizumab, nibolumab
PD-1
atezolizumab, avelumab, durvalumab
PD-L1
Targeted therapies
EGFR
gefitinib, erlotinib, afatinib, osimertinib
VEGFR
bevacizumab
ROS1
cizotinib,certinib
Inhibit
Specific cellular pathway/s other than proliferation

1.4.1 Structure and function of EGFR

EGFR is a surface protein expressed on actively growing epithelial cells (for example: hair follicles and skin cells) that belongs to a class of receptor tyrosine kinases known as ErbB1 or HER1. The EGFR protein is composed of distinct regions, an extracellular domain (N-termini), a hydrophobic transmembrane segment, and the intracellular domain (containing the kinase and C-termini) (61,62). EGFR on the cell surface exists as an inactive monomer. Activation of the EGFR pathway occurs when the extracellular domain of EGFR physically interacts with its ligand, Epidermal Growth Factor (EGF) leading to dimerization of EGFR and auto-phosphorylation of key tyrosine residues in the C-terminal by the kinase domain (61,62). The phosphorylated sites act as docking sites for adapter proteins containing Src-homology 2 domain (SH2) that in turn active a cascade of downstream signaling pathways, the EGFR signal transduction. The main downstream effectors of EGFR signal transduction includes the Ras/Raf/MAPK/ERK pathway and the STAT3/5 pathways resulting in cell proliferation, and the PI3K/Akt/mTOR pathways that facilitates cell survival and evasion from apoptosis (62,63) (**Figure 1.7**).

1.4.2 EGFR overexpression in NSCLC

Since EGFR is overexpressed in 15-35% of NSCLC patients (**Figure 1.5**) (27,31,40), and EGFR serves as a regulator of proliferation, the role of EGFR as a driver of NSCLC has been extensively studied. Apart from upregulation of wildtype EGFR leading to increased EGFR signaling, a few mutations such as amino acid substitution in exon 21 (L858R) or an in-frame deletion in exon 19, which collectively account for >90% of EGFR mutations in NSCLC lead to constitutive EGFR active (64–66). Such mutations are regarded as activating mutations because the kinase domain remains constitutively activated, which in a wildtype EGFR monomer in the absence of EGF remains autoinhibited preventing spurious signaling.

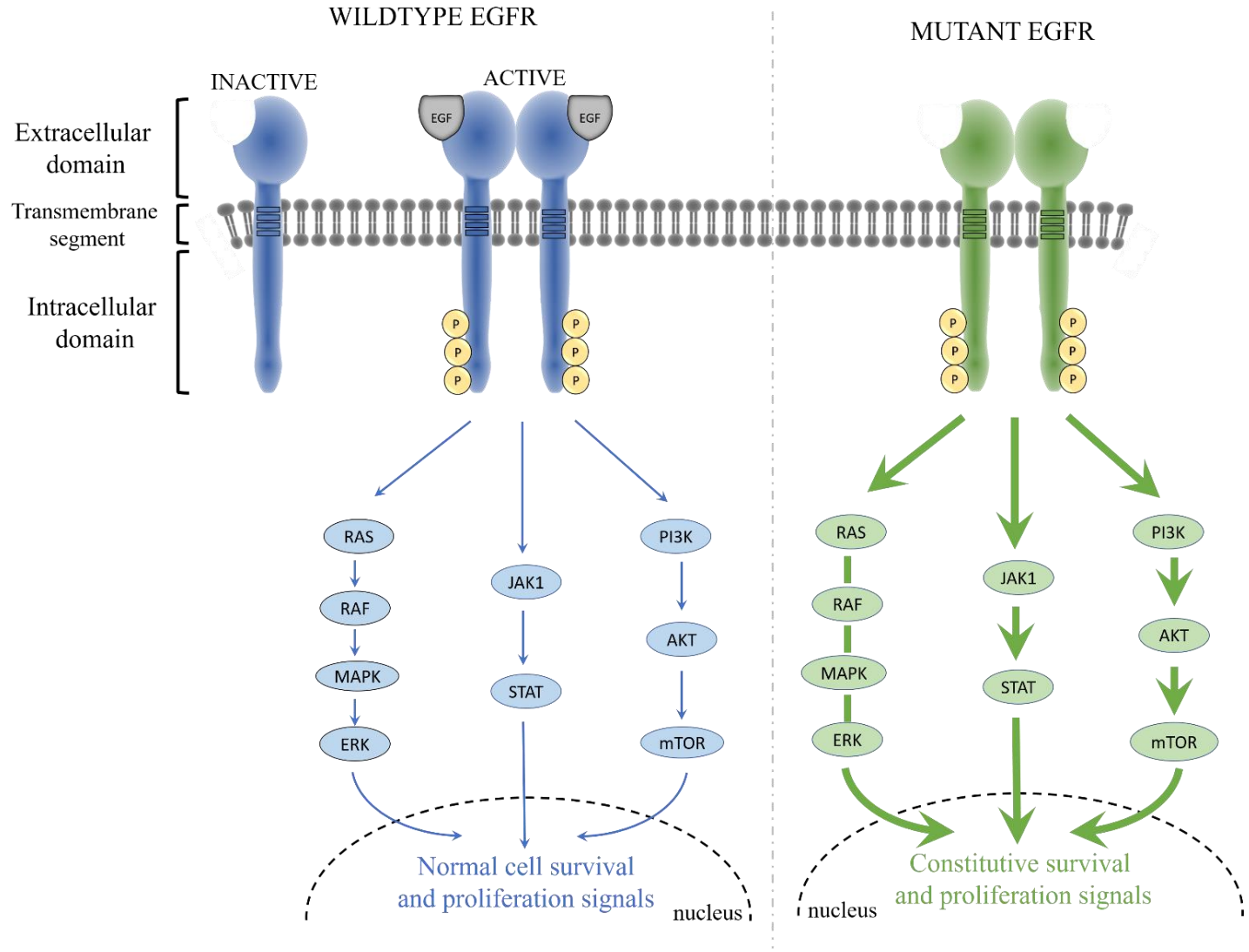
1.5 EGFR-Tyrosine Kinase Inhibitors

Patients diagnosed with an activating EGFR mutation are administered targeted therapies specifically designed to inhibit constitutively active mutant positive EGFR mediated signal transduction (**Figure 1.7**). Such inhibitors are called EGFR-Tyrosine Kinase Inhibitors (EGFR-TKIs). EGFR-TKIs have enhanced affinity for mutated EGFR molecules relative to wildtype

EGFR (67,68). EGFR-TKIs act as competitive inhibitors, competing with Adenosine-Triphosphate (ATP) for the ATP binding site of the constitutively active kinase domain, consequently blocking downstream signaling pathways (67,69). Although EGFR-TKIs display selectivity for EGFR-mutation positive tumors, EGFR-TKIs have also displayed an overall survival benefit in patients with tumors overexpressing wildtype EGFR. In such patients, EGFR-TKIs are commonly used as a first-line therapy, or as a second-line therapy after conventional chemotherapies have failed to clinically benefit the patient (70–73). Overall, EGFR-TKIs function by abrogating EGFR-signaling resulting in growth inhibition of the EGFR dependent cancer cells by overcoming "oncogene addiction". Based on the targeted mechanism of action of EGFR-TKIs, currently three generations of TKIs have been approved by the Food and Drug Administration (FDA) that are used as standard-of-care for NSCLC patients.

The first generation EGFR-TKIs include gefitinib and erlotinib that benefitted EGFR mutation positive patients in terms of overall and progression free survival (64,66,74). However, as the first generation of EGFR-TKIs showed reversible binding kinetics, and iterative competition with ATP for kinase domain binding, a second generation of TKIs were designed that irreversibly inhibit EGFR activity, including afatinib (75,76). Unfortunately, patients administered with either first or second generation EGFR-TKIs suffer with relapse of the disease due to development of resistance mediated by a common secondary mutation, i.e. an amino acid substitution in exon 20, T790M, described in the following section in detail (64,66,74–76). Therefore, to combat resistance mediated by T790M, a third generation of EGFR-TKIs was generated, osimertinib. Although osimertinib overcomes T790M mutation, patients treated with this drug suffer another secondary point mutation, C797S that causes resistance to osimertinib (77,78).

Figure 1.7. EGFR structure, and signal transduction mediated by wildtype and mutant EGFR. A monomer of EGFR is inactive, until it binds to its ligand, EGF that leads to dimerization with another EGFR monomer bound to EGF, followed by autophosphorylation of tyrosine residues in the intracellular domain. This triggers a cascade of signaling pathways in a normal cell. However, in NSCLC tumors with a mutated EGFR, EGFR signal transduction is constitutively active resulting in uncontrolled proliferation and constant survival signals (34–36).


WILDTYPE EGFR
MUTANT EGFR
INACTIVE
ACTIVE
EGF
EGF
Extracellular domain
Transmembrane segment
Intracellular domain
P
RAS
RAF
MAPK
ERK
JAK1
STAT
PI3K
AKT
mTOR
Normal cell survival and proliferation signals
nucleus
Constitutive survival and proliferation signals
nucleus

Despite improved clinical outcomes post EGFR-TKI treatment for NSCLC patients, patients suffer relapse of disease due to development of acquired resistance. As the EGFR-TKI of interest in this study is erlotinib, the following section will describe the development of acquired resistance to erlotinib.

1.5.1 Development of acquired erlotinib resistance

NSCLC tumors regress rapidly when patients are treated with erlotinib, however, within a year post-treatment, the majority of patients develop resistance. This is currently the major drawback of using erlotinib as a standard-of-care. Molecularly, resistance to erlotinib occurs when the drug exposure is inadequate at targeting active EGFR or the cell no longer depends on EGFR activity for survival, and the cancer progresses. Therefore, genomic profiling of tumors and cell-free DNA from patients post-erlotinib treatment reveal that erlotinib resistance in NSCLC cells is mediated via two major mechanisms (68,79–84) (**Figure 1.8**): 1) acquisition of secondary mutations that change the conformation of EGFR rendering it insensitive to erlotinib, 2) activation of alternate mechanisms of growth and proliferation, referred to as bypass tracks (80,85).

1.5.1.1 Resistance mediated via alterations in EGFR target:

Genetic testing of erlotinib treated NSCLC tumors reveal that over 50% of tumors incur a single recurrent secondary mutation in EGFR mutant tumors, substituting threonine at the 790th amino acid position in exon 20 with Methionine, i.e. T790M (68,86). Threonine at this position serves as a gatekeeper of the ATP-binding hydrophobic pocket. Substitution of the threonine with the bulky methionine causes a conformational change resulting in steric hinderance for erlotinib binding, but increases affinity for ATP (68,86). The resultant effect is constitutive activation of mutant EGFR and escape from erlotinib mediated inhibition.

Moreover, in approximately 10% of the NSCLC cases wild-type EGFR amplification co-exists with T790M mutation (86). It remains to be identified if the increased wild-type EGFR activity enhances resistance mediated by T790M mutation by increasing the erlotinib requirement to inhibit total EGFR, or if it is a passenger event. In addition to T790M mutation, a few other *in vitro* tested mutations that have been reported to cause acquired erlotinib resistance in NSCLC

patients include D761Y (87,88), L747S (89) and T854A (90). However, these mutations in EGFR comprise only 1-2 % of erlotinib resistant NSCLC cases.

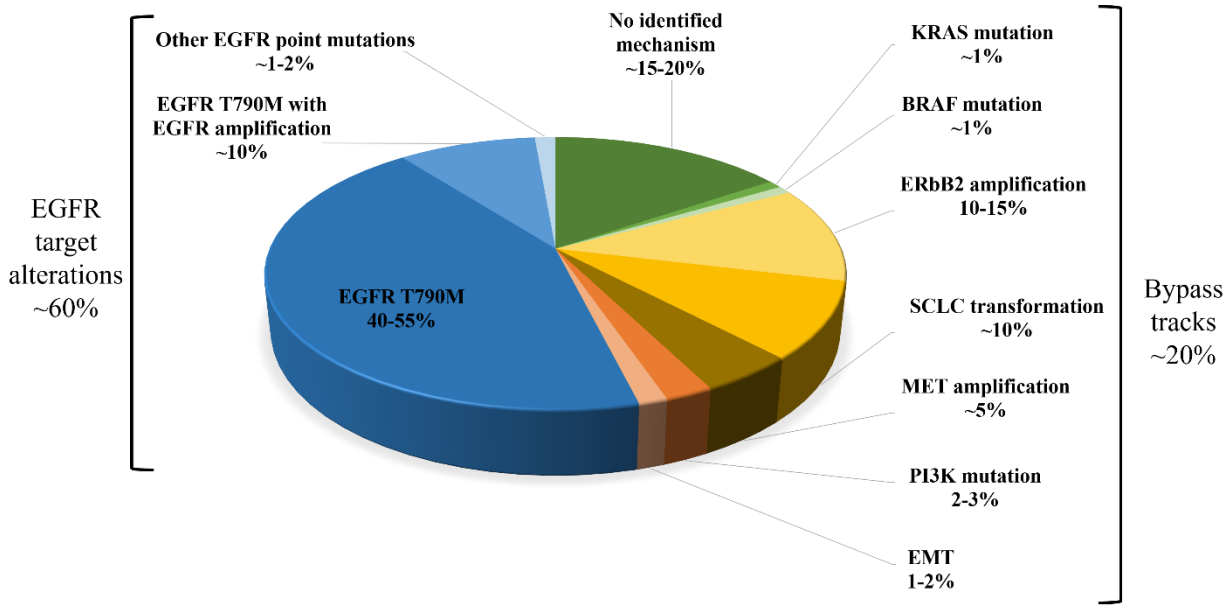


Figure 1.8. Mechanisms of development of acquired erlotinib resistance. Erlotinib resistance in NSCLC is either mediated by alterations made to the target protein, EGFR or via activation of bypass tracks or alternative pathways (80,85).

1.5.1.2 Erlotinib resistance due to activation of bypass tracks:

Post-treatment with erlotinib, approximately 20% of NSCLC patients develop resistance through the activation of bypass tracks or alternative pathways (80,85). Following activation of bypass tracks, the effect of erlotinib treatment is insufficient to prevent tumor growth. These resistant cells bypass their dependence on EGFR pathway via activation of alternative mechanisms that trigger the same key downstream growth and survival signaling pathways such as PI3K, MAPK/ERK and Akt (**Figure 1.7**). Such bypass tracks that confer EGFR-independent growth advantage in the presence of erlotinib include: 1) activation of oncogenic pathways driven by MET, BRAF, HER2, HER3, PIK3CA, IGF1R, FGF1, 2) loss of function of tumor-suppressor genes such as PTEN and NF1, or 3) acquired ability of histological transformation, i.e. transform into small cell lung cancer (SCLC) or epithelial to mesenchymal transition (EMT) (80–84). A few major bypass mechanisms from **Figure 1.8** are described below:

a) MET amplification: One of the first bypass tracks to be described in TKI resistance in EGFR-mutant NSCLC was MET amplification (91). The gene that encodes the proto-oncogene MET was amplified in 22% of tumors tested in the study, whereas downregulation of MET re-sensitized the tumors to the TKI gefitinib. Moreover, it was identified that MET amplification triggered the PI3K pathway that enhanced cell survival (91). Therefore, since gefitinib and erlotinib are both first generation TKIs, with a common chemical backbone and similar clinical efficacies (92–94), MET has also been studied extensively as a mechanism of acquired resistance to erlotinib (84,95). In addition to activation of survival pathways, MET also mediates resistance via EMT that results in histological transformation, rendering erlotinib ineffective (84). Currently, a clinical trial to evaluate the efficacy of the MET inhibitor, capmatinib in combination with erlotinib is ongoing (96).
b) ErbB2/HER2 amplification: One of the most common bypass mechanisms of erlotinib resistance is amplification of HER2, which also belongs to the ErbB2 family of receptor proteins, the same family as EGFR (also known as ErbB1). It is well documented that EGF can trigger EGFR signaling when EGFR forms heterodimers with other members of the ErbB family of proteins (94,97). Therefore, it is not surprising that 10-15% of erlotinib treated patients develop resistance via HER2 overexpression which compensates for the inhibited EGFR activity (80,85) (**Figure 1.8**). Currently, there are no HER2 targeting inhibitors in clinical trials that can be used to benefit patients with erlotinib resistant NSCLC tumors displaying HER2 amplification.
c) Histological transformation of NSCLC cells: Transformation of NSCLC cells is a mechanism that promotes erlotinib resistance by acquiring phenotypic plasticity to evade therapy. Phenotypic plasticity is described as a mechanism by which a cell undergoes reprogramming to express genes and phenotypes of another cell type that maybe intrinsically resistant to the therapy (98,99). Currently, approximately 10% of NSCLC cases that stop responding to erlotinib have been transformed from NSCLC into small cell lung cancer (SCLC) (80,85) (**Figure 1.3, 1.8**). Such transformed NSCLC cells display characteristics typical of SCLC tumors i.e. loss of functions of genes such as retinoblastoma (RB) and TP53 (81,100). Although still harboring the original EGFR mutation, these transformed erlotinib resistant tumors show dramatic loss of EGFR activity,

using alternative mechanisms to sustain their growth and survival (81,82). In addition to transformation from one lung cell type to another, 1-2% of NSCLC cases display EMT as a mechanism to evade therapy as well (80,85) (**Figure 1.8**). Upon transitioning from epithelial to mesenchymal states, epithelial cells lose their cell-to-cell junctions and acquire invasive characteristics. EMT is a hallmark of cancer cells that metastasize to distant organs, and are very aggressive (99,101). Therefore, acquiring EMT characteristics is regarded as a mechanism by which NSCLC cells acquire resistance to erlotinib (84).

1.6 Epigenetics in lung cancer:

As outlined above, the determinants of cancers, including NSCLC and the biomarkers associated with hallmarks of cancer have primarily been identified as genetic perturbations. However, with the shifting paradigm since the 1990s, novel epigenetic alterations that promote cancer and those associated with the various hallmarks of cancer are being actively investigated (102,103). Epigenetic alterations result in changes in gene expression without any alterations to the DNA sequence, and are somatically heritable (102,103). Various epigenetic alterations that promote tumorigenesis include abnormal DNA methylation patterns of promoters of genes involved in cancers, or disruption in posttranslational modification (PTM) of histones resulting in reorganization of the chromatin (102–104). Maintenance of a normal composition of the chromatin or the DNA methylation pattern are attributed to multiple epigenetic factors and several non-coding RNA regulators, many of which are also reported to be severely dysregulated in cancers, including lung cancer (102,103).

1.6.1 Role of DNA methylation in the development of lung cancers:

Hypomethylation of regions of the genome such as the promoters of genes result in epigenetic activation of the gene, observed as a mechanism of oncogene activation in lung cancers. For instance, the promoter of a prometastatic oncogenic protein, Transmembrane Serine Protease 4, TMPRSS4 is aberrantly hypomethylated in NSCLC and is associated with poor prognosis (105). Whereas in the case of tumor suppressors which often undergo loss of function during progression of NSCLC, aside from mutation facilitated silencing, are also reported to be inactivated due to hypermethylation of their promoters. For example, the tumor suppressor *RAS effector homolog*,

RASSF1A gene is hypermethylated in multiple cancers, including 30-50% of NSCLC cases (106–108) as is the tumor suppressive microRNA, miR-34a (109,110).

1.6.2 Role of histone modifications in the development of lung cancers:

Histone PTMs that are responsible for chromatin maintenance are often disrupted in cancers as well (102,103). Among others, a commonly observed preneoplastic change in lung and multiple other cancers is the global loss of a repressive PTM on histone 4 (H4), trimethylation of lysine 20 (H4K20me3)(4,111–113). As a result, genome integrity is disrupted resulting in increased expression of genes, specifically oncogenes, and defects in the DNA Damage Response (DDR) pathway (114–116).

1.6.3 Epigenetic modifiers as determinants of NSCLC, and mediators of TKI resistance:

In cancers, alterations in methylation status and histone modifications are the result of dysregulated epigenetic machinery, including 1) enzymes responsible for methylation or demethylation of DNA, 2) machinery involved in establishing histone PTM, and readers and erasers of the PTM, 3) enzymes responsible for histone remodeling, and 3) non-coding RNAs that function as epigenetic modifiers. Mutations leading to the loss of the DNA methyltransferase DNMT1, a key enzyme responsible for global methylation of the genome contributes to tumorigenesis through multiple mechanisms, including alterations in genomic stability and activation of multiple oncogenes (102–104,117,118). Similarly, loss of histone modifiers such as SUV420H2, the mediator of the repressive H4K20me3 modification functioning as a tumor suppressor, is also associated with poor prognosis in cancers (114,115). Conversely, overexpression of EHMT2, a histone methyltransferase that catalyzes mono- and dimethylation at histone H3 lysine 9 resulting in gene silencing, is commonly observed in multiple cancers(119). In erlotinib resistant NSCLC cells, EHMT2 is upregulated resulting in transcriptional abrogation of the tumor suppressor gene, *PTEN* (120).

MicroRNAs, small non-coding RNAs are also reported to contribute to the development of cancers and resistance to drugs such as TKIs. MicroRNAs function by targeting transcripts of genes, negatively regulate their expressions. In cells, microRNAs function as both post-translational regulators of target transcripts and as epigenetic modulators. A microRNA is regarded as an

epigenetic factor if it regulates the transcription of an epigenetic enzyme, or if the transcription of a microRNA is regulated via the function of an epigenetic modulator (103). In the case of tumor-suppressive microRNAs, miR-200b, miR-200c, and miR-205 are silenced in a sustained manner in NSCLC via sequential epigenetic modifications. Firstly, H3K27me3, a repressive PTM, catalyzed by the pro-tumorigenic epigenetic factor, EZH2 results in downregulation of the microRNAs miR-200b, c and -205, followed by severe hypermethylation of their respective promoters (103,121,122). A distinct mode of epigenetic repression for a microRNA is found for the *mir-34a* gene. Epigenetic promoter methylation of *mir-34a* in NSCLC results in loss of miR-34a, a prominent p53-regulated tumor suppressive microRNA(109,110). Such tumor suppressive microRNAs are negatively regulated, indirectly mediating tumorigenesis due to epigenetic remodeling of the genome. MicroRNAs also serve as epigenetic mediators of erlotinib resistance. One such microRNA, miR-17-5p which contextually functions as a tumor suppressive microRNA is downregulated in erlotinib resistant NSCLC cells, resulting in upregulation of the pro-tumorigenic epigenetic remodeler, EZH1(122,123).

Although recent investigations have identified roles for a few epigenetic factors as mediators of erlotinib resistance in NSCLC, with the growing knowledge of epigenetic modifiers and the increasing number of microRNAs, it can be speculated that more such novel mediators are yet to be identified.

1.7 Significance and goals of this study:

Despite the identification of multiple signaling pathways that facilitate erlotinib resistance, in 15-20% of cases that have acquired resistance, the mechanisms remain to be identified (124,125) and in many cases the molecular contributions involved in regulating the bypass tracks is unknown (**Figure 1.8**). Histologic analysis and genetic screening are incapable of identifying non-genetic mechanisms that may have a role in development of resistance, such as alterations in epigenetic mediators or microRNAs. To this end, the aim of this study is to identify novel mediators of erlotinib resistance. To address this, two simultaneous screens were conducted in erlotinib sensitive cells: 1) A genome-wide knock-out screen using the CRISPR-Cas9 system to identify both protein coding and microRNA genes that when lost lead to erlotinib resistance, and 2) a microRNA overexpression screen where a library of individually arrayed microRNAs are

exogenously transfected in the cells to identify novel microRNAs that when overexpressed lead to erlotinib resistance.

1.8 References:

1. Pal AS, Kasinski AL. Animal Models to Study MicroRNA Function. Adv Cancer Res. Academic Press Inc.; 2017. page 53–118.

2. Cadieux BB, Ching T-T, Vandenberg SR, Costello JF. Genome-wide Hypomethylation in Human Glioblastomas Associated with Specific Copy Number Alteration, Methylenetetrahydrofolate Reductase Allele Status, and Increased Proliferation. Cancer Res [Internet]. 2006 [cited 2020 Jan 17];66:8469–76. Available from: http://cancerres.aacrjournals.org/

3. Celeste A, Difilippantonio S, Difilippantonio MJ, Fernandez-Capetillo O, Pilch DR, Sedelnikova OA, et al. H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. Cell. Cell Press; 2003;114:371–83.

4. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callén E, et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev. Cold Spring Harbor Laboratory Press; 2008;22:2048–61.

5. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell. 2011. page 646–74.

6. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature. Nature Publishing Group; 2013;502:333–9.

7. Kemp R, Godwin AK, Prados M, Zwarthoff EC, Disaia P, Evason K, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;174:1034–5.

8. MURPHY KL, DENNIS AP, ROSEN JM. A gain of function p53 mutant promotes both genomic instability and cell survival in a novel p53-null mammary epithelial cell model. FASEB J [Internet]. 2000 [cited 2020 Jan 17];14:2291–302. Available from: http://www.fasebj.org/doi/10.1096/fj.00-0128com

9. Mantovani F, Collavin L, Del Sal G. Mutant p53 as a guardian of the cancer cell. Cell Death Differ. Nature Publishing Group; 2019. page 199–212.

10. Willis A, Jung EJ, Wakefield T, Chen X. Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. Oncogene. 2004;23:2330–8.

11. Luo B, Wing H, Subramanian A, Sharifnia T, Okamoto M, Yang X, et al. Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci. 2008;105:20380–20385.

12. Liu P, Wang Y, Li X. Targeting the untargetable KRAS in cancer therapy. Acta Pharm. Sin. B. Chinese Academy of Medical Sciences; 2019. page 871–9.

13. Yang H, Liang SQ, Schmid RA, Peng RW. New horizons in KRAS-mutant lung cancer: Dawn after darkness. Front. Oncol. Frontiers Media S.A.; 2019.

14. Singh A, Greninger P, Rhodes D, Koopman L, Violette S, Bardeesy N, et al. A Gene Expression Signature Associated with "K-Ras Addiction" Reveals Regulators of EMT and Tumor Cell Survival. Cancer Cell. 2009;15:489–500.

15. Arbour KC, Jordan E, Kim HR, Dienstag J, Yu HA, Sanchez-Vega F, et al. Effects of co-occurring genomic alterations on outcomes in patients with KRAS-mutant non-small cell lung cancer. Clin Cancer Res. American Association for Cancer Research Inc.; 2018;24:334–40.

16. Collins FS. The Cancer Genome Atlas (TCGA). Online. 2007;1–17.

17. Zhang L, Zhang S, Yao J, Lowery FJ, Zhang Q, Huang WC, et al. Microenvironment-induced PTEN loss by exosomal microRNA primes brain metastasis outgrowth. Nature [Internet]. 2015;527:100–4. Available from: http://dx.doi.org/10.1038/nature15376

18. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat. Med. 2013. page 1423–37.

19. Flister MJ, Bergom C. Genetic Modifiers of the Breast Tumor Microenvironment. Trends in Cancer. Cell Press; 2018. page 429–44.

20. Carvalho PD, Guimarães CF, Cardoso AP, Mendonça S, Costa ÂM, Oliveira MJ, et al. KRAS oncogenic signaling extends beyond cancer cells to orchestrate the microenvironment. Cancer Res. American Association for Cancer Research Inc.; 2018. page 7–14.

21. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell. Cell Press; 2016;165:35–44.

22. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. Immunity. Cell Press; 2018;48:812-830.e14.

23. Raphael BJ, Dobson JR, Oesper L, Vandin F. Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine. Genome Med [Internet]. 2014 [cited 2020 Jan 31];6:5. Available from: http://genomemedicine.biomedcentral.com/articles/10.1186/gm524

24. Marsit CJ, Zheng S, Aldape K, Hinds PW, Nelson HH, Wiencke JK, et al. PTEN expression in non-small-cell lung cancer: Evaluating its relation to tumor characteristics, allelic loss, and epigenetic alteration. Hum Pathol. 2005;36:768–76.

25. Sos ML, Koker M, Weir BA, Heynck S, Rabinovsky R, Zander T, et al. PTEN loss contributes to erlotinib resistance in EGFR-mutant lung cancer by activation of akt and EGFR. Cancer Res. 2009;69:3256–61.

26. Zhang M, Zhuang G, Sun X, Shen Y, Wang W, Li Q, et al. TP53 mutation-mediated genomic instability induces the evolution of chemoresistance and recurrence in epithelial ovarian cancer.

27. Lung Cancer 101 | LUNGevity Foundation [Internet]. [cited 2020 Jan 19]. Available from: https://lungevity.org/for-patients-caregivers/lung-cancer-101

28. Lung Cancer 101 | Lungcancer.org [Internet]. [cited 2020 Jan 18]. Available from: https://www.lungcancer.org/find_information/publications/163-lung_cancer_101/269-non_small_cell_lung_cancer_treatment

29. Lung cancer atlas. The Lung Cancer https://canceratlas.cancer.org/the-burden/lung-cancer/ [Internet]. [cited 2019 Dec 22]. Available from: https://canceratlas.cancer.org/the-burden/lung-cancer/

30. Lung Cancer Screening [Internet]. [cited 2020 Jan 19]. Available from: https://www.radiologyinfo.org/en/info.cfm?pg=screening-lung

31. Schabath MB, Cote ML. Cancer progress and priorities: Lung cancer. Cancer Epidemiol Biomarkers Prev. 2019;28:1563–79.

32. Breeze R, Turk M. Cellular structure, function and organization in the lower respiratory tract. Environ Health Perspect. 1984;VOL. 55:3–24.

33. Chansky K, Detterbeck FC, Nicholson AG, Rusch VW, Vallières E, Groome P, et al. The IASLC Lung Cancer Staging Project: External Validation of the Revision of the TNM Stage Groupings in the Eighth Edition of the TNM Classification of Lung Cancer. J Thorac Oncol. Elsevier Inc; 2017;12:1109–21.

34. Park KS, Liang MC, Raiser DM, Zamponi R, Roach RR, Curtis SJ, et al. Characterization of the cell of origin for small cell lung cancer. Cell Cycle. 2011;10:2806–15.

35. Nakazawa K, Kurishima K, Tamura T, Kagohashi K, Ishikawa H, Hiroaki S, et al. Specific organ metastases and survival in small cell lung cancer. Oncol Lett. 2012;4:617–20.

36. The Dominant Malignancy: Lung Cancer : Of Schemes and Memes Blog [Internet]. [cited 2020 Jan 18]. Available from: http://blogs.nature.com/ofschemesandmemes/2014/09/11/the-dominant-malignancy-lung-cancer

37. Hynds RE, Janes SM. Airway basal cell heterogeneity and lung squamous cell carcinoma. Cancer Prev. Res. American Association for Cancer Research Inc.; 2017. page 491–3.

38. Wistuba II, Gazdar AF. Molecular Biology of Preneoplastic Lesions of the Lung. Lung Cancer, Third Ed. Blackwell Publishing Ltd; 2008. page 84–98.

39. Understanding Your Pathology Report: Lung Cancer In Situ [Internet]. [cited 2020 Mar 7]. Available from: https://www.cancer.org/treatment/understanding-your-diagnosis/tests/understanding-your-pathology-report/lung-pathology/lung-cancer-in-situ-pathology.html

40. Non-Small Cell Lung Carcinoma - My Cancer Genome [Internet]. [cited 2020 Jan 18]. Available from: https://www.mycancergenome.org/content/disease/non-small-cell-lung-carcinoma/

41. Hirsch FR, Varella-Garcia M, Bunn PA, Di Maria M V., Veve R, Bremnes RM, et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: Correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol. 2003;21:3798–807.

42. Onitsuka T, Uramoto H, Ono K, Takenoyama M, Hanagiri T, Oyama T, et al. Comprehensive molecular analyses of lung adenocarcinoma with regard to the epidermal growth factor Receptor, K-ras, MET, and hepatocyte growth factor status. J Thorac Oncol. Lippincott Williams and Wilkins; 2010;5:591–6.

43. Howe LR, Leevers SJ, Gómez N, Nakielny S, Cohen P, Marshall CJ. Activation of the MAP kinase pathway by the protein kinase raf. Cell [Internet]. 1992 [cited 2020 Feb 2];71:335–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1330321

44. Riely GJ, Marks J, Pao W. KRAS mutations in non-small cell lung cancer. Proc Am Thorac Soc. 2009. page 201–5.

45. Lee JW, Soung YH, Kim SY, Nam SW, Park WS, Wang YP, et al. ERBB2 kinase domain mutation in the lung squamous cell carcinoma. Cancer Lett. 2006;237:89–94.

46. Vijayalakshmi R, Krishnamurthy A. Targetable "Driver" Mutations in Non Small Cell Lung Cancer. Indian J. Surg. Oncol. 2011. page 178–88.

47. Luo SY, Lam DC. Oncogenic driver mutations in lung cancer. Transl Respir Med. Springer Nature; 2013;1.

48. An SJ, Chen ZH, Su J, Zhang XC, Zhong WZ, Yang JJ, et al. Identification of enriched driver gene alterations in subgroups of non-small cell lung cancer patients based on histology and smoking status. PLoS One. 2012;7.

49. Frampton GM, Ali SM, Rosenzweig M, Chmielecki J, Lu X, Bauer TM, et al. Activation of MET via Diverse Exon 14 Splicing Alterations Occurs in Multiple Tumor Types and Confers Clinical Sensitivity to MET Inhibitors. Michiana Hematol [Internet]. Thoracic Oncology Program; 2015 [cited 2020 Jan 19]; Available from: www.aacrjournals.org

50. Awad MM, Oxnard GR, Jackman DM, Savukoski DO, Hall D, Shivdasani P, et al. MET Exon 14 Mutations in Non-Small-Cell Lung Cancer Are Associated With Advanced Age and Stage-Dependent MET Genomic Amplification and c-Met Overexpression. J Clin Oncol [Internet]. 2016 [cited 2020 Jan 19];34:721–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26729443

51. MET Dysregulation | Understanding MET in NSCLC [Internet]. [cited 2020 Jan 19]. Available from: https://www.seemetgene.com/met-dysregulation-non-small-cell-lung-cancer/?site=CTB-1221016GS100001&source=01030&gclid=Cj0KCQiAgKzwBRCjARIsABBbFuhFQBbrkKDwi4ZH86MIuu8TdjdatQ-P30uXVME8gtQhJh0oH0h7T1AaAihQEALw_wcB&gclsrc=aw.ds

52. Soria J-C, Lee H-Y, Lee JI, Wang L, Issa J-P, Kemp BL, et al. Lack of PTEN Expression in Non-Small Cell Lung Cancer Could Be Related to Promoter Methylation 1. 2002.

53. Lee SY, Kim MJ, Jin G, Yoo SS, Park JY, Choi JE, et al. Somatic mutations in epidermal growth factor receptor signaling pathway genes in non-small cell lung cancers. J Thorac Oncol. Lippincott Williams and Wilkins; 2010;5:1734–40.

54. Zysman MA, Chapman WB, Bapat B. Considerations when analyzing the methylation status of PTEN tumor suppressor gene. Am J Pathol. American Society for Investigative Pathology Inc.; 2002;160:795–800.

55. Howington JA, Blum MG, Chang AC, Balekian AA, Murthy SC. Treatment of stage I and II non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American college of chest physicians evidence-based clinical practice guidelines. Chest. 2013;143.

56. Non-Small Cell Lung Cancer Treatment (PDQ®)–Health Professional Version - National Cancer Institute [Internet]. [cited 2020 Jan 19]. Available from: https://www.cancer.gov/types/lung/hp/non-small-cell-lung-treatment-pdq#link/_485420

57. Mazzone P, Mekhail T. Current and emerging medical treatments for non-small cell lung cancer: A primer for pulmonologists. Respir. Med. 2012. page 473–92.

58. Kris MG, Natale RB, Herbst RS, Lynch TJ, Prager D, Belani CP, et al. Efficacy of Gefitinib, an Inhibitor of the Epidermal Growth Factor Receptor Tyrosine Kinase, in Symptomatic Patients with Non-Small Cell Lung Cancer: A Randomized Trial. J Am Med Assoc. 2003;290:2149–58.

59. Fukuoka M, Yano S, Giaccone G, Tamura T, Nakagawa K, Douillard JY, et al. Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. J Clin Oncol. 2003;21:2237–46.

60. Hanna N, Johnson D, Temin S, Baker S, Brahmer J, Ellis PM, et al. Systemic Therapy for Stage IV Non-Small-Cell Lung Cancer: American Society of Clinical Oncology Clinical Practice Guideline Update. J Clin Oncol [Internet]. 2017 [cited 2020 Jan 19];35:3484–515. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28806116

61. Jura N, Endres NF, Engel K, Deindl S, Das R, Lamers MH, et al. Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. Cell. Cell Press; 2009;137:1293–307.

62. Sibilia M, Kroismayr R, Lichtenberger BM, Natarajan A, Hecking M, Holcmann M. The epidermal growth factor receptor: From development to tumorigenesis. Differentiation. 2007;75:770–87.

63. Pao W, Riely GJ, de Stanchina E, Aoe K, Shien K, Miller VA, et al. Lung cancers with acquired resistance to EGFR inhibitors occasionally harbor BRAF gene mutations but lack mutations in KRAS, NRAS, or MEK1. Proc Natl Acad Sci. 2012;109:E2127–33.

64. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung, cancer: Correlation with clinical response to gefitinib therapy. Science (80-). 2004;304:1497–500.

65. Sharma S V, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer [Internet]. 2007;7:169–81. Available from: http://www.nature.com/nrc/journal/v7/n3/full/nrc2088.html

66. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med [Internet]. 2004;350:2129–39. Available from: http://dx.doi.org/10.1056/NEJMoa040938

67. Carey KD, Garton AJ, Romero MS, Kahler J, Thomson S, Ross S, et al. Kinetic analysis of epidermal growth factor receptor somatic mutant proteins shows increased sensitivity to the epidermal growth factor receptor tyrosine kinase inhibitor, erlotinib. Cancer Res. 2006;66:8163–71.

68. Yun CH, Mengwasser KE, Toms A V., Woo MS, Greulich H, Wong KK, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc Natl Acad Sci U S A. 2008;105:2070–5.

69. Yun CH, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, et al. Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. Cancer Cell. 2007;11:217–27.

70. Osarogiagbon RU, Cappuzzo F, Ciuleanu T, Leon L, Klughammer B. Erlotinib therapy after initial platinum doublet therapy in patients with EGFR wild type non-small cell lung cancer: Results of a combined patient-level analysis of the NCIC CTG BR.21 and SATURN trials. Transl Lung Cancer Res. AME Publishing Company; 2015;4:465–74.

71. Hirai F, Edagawa M, Shimamatsu S, Toyozawa R, Toyokawa G, Nosaki K, et al. Evaluation of erlotinib for the treatment of patients with non-small cell lung cancer with epidermal growth factor receptor wild type. Oncol Lett. Spandidos Publications; 2017;14:306–12.

72. Inno A, Di Noia V, Martini M, D'Argento E, Di Salvatore M, Arena V, et al. Erlotinib for Patients with EGFR Wild-Type Metastatic NSCLC: a Retrospective Biomarkers Analysis. Pathol Oncol Res. Springer Netherlands; 2019;25:513–20.

73. Pérol M, Chouaid C, Pérol D, Barlési F, Gervais R, Westeel V, et al. Randomized, phase III study of gemcitabine or erlotinib maintenance therapy versus observation,with predefined second-line treatment, after cisplatin-gemcitabine induction chemotherapy in advanced non-small-cell lung cancer. J Clin Oncol. 2012;30:3516–24.

74. Bonomi PD, Buckingham L, Coon J. Selecting Patients for Treatment with Epidermal Growth Factor Tyrosine Kinase Inhibitors. Clin Cancer Res [Internet]. 2007 [cited 2020 Jan 20];13. Available from: www.aacrjournals.org

75. Joshi M, Rizvi SM, Belani CP. Afatinib for the treatment of metastatic non-small cell lung cancer. Cancer Manag Res. 2015;7:75–82.

76. Wu SG, Liu YN, Tsai MF, Chang YL, Yu CJ, Yang PC, et al. The mechanism of acquired resistance to irreversible EGFR tyrosine kinase inhibitor-afatinib in lung adenocarcinoma patients. Oncotarget. 2016;7:12404–13.

77. Madic J, Jovelet C, Lopez J, André B, Fatien J, Miran I, et al. EGFR C797S, EGFR T790M and EGFR sensitizing mutations in non-small cell lung cancer revealed by six-color crystal digital PCR [Internet]. Available from: www.oncotarget.com

78. Wang S, Cang S, Liu D. Third-generation inhibitors targeting EGFR T790M mutation in advanced non-small cell lung cancer. J Hematol Oncol [Internet]. Journal of Hematology & Oncology; 2016;9:1–7. Available from: http://dx.doi.org/10.1186/s13045-016-0268-z

79. Demuth C, Madsen AT, Weber B, Wu L, Meldgaard P, Sorensen BS. The T790M resistance mutation in EGFR is only found in cfDNA from erlotinib-treated NSCLC patients that harbored an activating EGFR mutation before treatment. BMC Cancer [Internet]. BioMed Central; 2018 [cited 2019 Oct 13];18:191. Available from: https://bmccancer.biomedcentral.com/articles/10.1186/s12885-018-4108-0

80. Tetsu O, Hangauer MJ, Phuchareon J, Eisele DW, McCormick F. Drug Resistance to EGFR Inhibitors in Lung Cancer. Chemotherapy. 2016;61:223–35.

81. Niederst MJ, Sequist L V., Poirier JT, Mermel CH, Lockerman EL, Garcia AR, et al. RB loss in resistant EGFR mutant lung adenocarcinomas that transform to small-cell lung cancer. Nat Commun. Nature Publishing Group; 2015;6:6377.

82. Sequist L V., Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med. 2011;3:75ra26.

83. Shien K, Toyooka S, Yamamoto H, Soh J, Jida M, Thu KL, et al. Acquired resistance to EGFR inhibitors is associated with a manifestation of stem cell-like properties in cancer cells. Cancer Res. 2013;73:3051–61.

84. Jakobsen KR, Demuth C, Madsen AT, Hussmann D, Vad-Nielsen J, Nielsen AL, et al. MET amplification and epithelial-to-mesenchymal transition exist as parallel resistance mechanisms in erlotinib-resistant, EGFR-mutated, NSCLC HCC827 cells. Oncogenesis. Nature Publishing Group; 2017;6.

85. Camidge DR, Pao W, Sequist L V. Acquired resistance to TKIs in solid tumours: learning from lung cancer. Nat Rev Clin Oncol [Internet]. 2014;11:473–81. Available from: http://dx.doi.org/10.1038/nrclinonc.2014.104

86. Ercan D, Zejnullahu K, Yonesaka K, Xiao Y, Capelletti M, Rogers A, et al. Amplification of EGFR T790M causes resistance to an irreversible EGFR inhibitor. Oncogene. 2010;29:2346–56.

87. Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, et al. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. Clin Cancer Res. 2006;12:6494–501.

88. Toyooka S, Date H, Uchida A, Kiura K, Takata M. The epidermal growth factor receptor D761Y mutation and effect of tyrosine kinase inhibitor [1]. Clin. Cancer Res. 2007. page 3431.

89. Costa DB, Schumer ST, Tenen DG, Kobayashi S. Differential responses to erlotinib in epidermal growth factor receptor (EGFR)-mutated lung cancers with acquired resistance to gefitinib carrying the L747S or T790M secondary mutations. J Clin Oncol [Internet]. 2008 [cited 2020 Jan 20];26:1182–4; author reply 1184-6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18309959

90. Bean J, Riely GJ, Balak M, Marks JL, Ladanyi M, Miller VA, et al. Acquired resistance to epidermal growth factor receptor kinase inhibitors associated with a novel T854A mutation in a patient with EGFR-mutant lung adenocarcinoma. Clin Cancer Res. 2008;14:7519–25.

91. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Joon OP, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science (80-). 2007;316:1039–43.

92. Speake G, Marshall G, Anderton J, Acheson K, Logie A, Vincent J, et al. A pharmacological comparison of gefitinib (IRESSA) and erlotinib. Cancer Res. 2006;66.

93. Perez-Soler R. The role of erlotinib (Tarceva, OSI 774) in the treatment of non-small cell lung cancer. Clin Cancer Res. 2004.

94. Xu Y, Liu H, Chen J, Zhou Q. Acquired resistance of lung adenocarcinoma to EGFR-tyrosine kinase inhibitors gefitinib and erlotinib. Cancer Biol Ther [Internet]. 2010 [cited 2020 Jan 20];9:572–82. Available from: https://doi.org/10.4161/cbt.9.8.11881

95. Politi K, Fan PD, Shen R, Zakowski M, Varmus H. Erlotinib resistance in mouse models of epidermal growth factor receptor-induced lung adenocarcinoma. Dis Model Mech. 2010;3:111–9.

96. Clinical Trial: NCT02468661 - My Cancer Genome [Internet]. [cited 2020 Jan 20]. Available from: https://www.mycancergenome.org/content/clinical_trials/NCT02468661/

97. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat. Rev. Mol. Cell Biol. 2001. page 127–37.

98. Hammerlindl H, Schaider H. Tumor cell-intrinsic phenotypic plasticity facilitates adaptive cellular reprogramming driving acquired drug resistance. J. Cell Commun. Signal. Springer Netherlands; 2018. page 133–41.

99. Emmons MF, Faião-Flores F, Smalley KSM. The role of phenotypic plasticity in the escape of cancer cells from targeted therapy. Biochem Pharmacol [Internet]. Elsevier Inc.; 2016;122:1–9. Available from: http://dx.doi.org/10.1016/j.bcp.2016.06.014

100. Lee JK, Lee J, Kim S, Kim S, Youk J, Park S, et al. ClonalHistory & genetic predictors of transformation into small-cell carcinomas from lung adenocarcinomas. J Clin Oncol. American Society of Clinical Oncology; 2017;35:3065–74.

101. Thiery JP, Acloque H, Huang RYJ, Nieto MA. Epithelial-Mesenchymal Transitions in Development and Disease. Cell. 2009. page 871–90.

102. Baylin SB, Jones PA. Epigenetic determinants of cancer. Cold Spring Harb Perspect Biol. Cold Spring Harbor Laboratory Press; 2016;8:a019505.

103. Quintanal-Villalonga Á, Molina-Pinelo S. Epigenetics of lung cancer: a translational perspective. Cell. Oncol. Springer; 2019. page 739–56.

104. Lin RK, Hsu HS, Chang JW, Chen CY, Chen JT, Wang YC. Alteration of DNA methyltransferases contributes to 5′CpG methylation and poor prognosis in lung cancer. Lung Cancer. Elsevier; 2007;55:205–13.

105. Villalba M, Diaz-Lagares A, Redrado M, De Aberasturi AL, Segura V, Bodegas ME, et al. Epigenetic alterations leading to TMPRSS4 promoter hypomethylation and protein overexpression predict poor prognosis in squamous lung cancer patients. Oncotarget. Impact Journals LLC; 2016;7:22752–69.

106. Marsit CJ, Kim DH, Liu M, Hinds PW, Wiencke JK, Nelson HH, et al. Hypermethylation of RASSF1A and BLU tumor suppressor genes in non-small cell lung cancer: Implications for tobacco smoking during adolescence. Int J Cancer [Internet]. 2005 [cited 2020 Mar 15];114:219–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15540210

107. Liu W-J, Tan X-H, Guo B-P, Ke Q, Sun J, Cen H. Associations between RASSF1A promoter methylation and NSCLC: a meta-analysis of published data. Asian Pac J Cancer Prev [Internet]. 2013 [cited 2020 Mar 15];14:3719–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23886171

108. Liu Y, Gao W, Siegfried JM, Weissfeld JL, Luketich JD, Keohavong P. Promoter methylation of RASSF1A and DAPK and mutations of K-ras, p53, and EGFR in lung tumors from smokers and never-smokers. BMC Cancer [Internet]. BioMed Central; 2007 [cited 2020 Mar 15];7:74. Available from: http://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-7-74

109. Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N, et al. Transcriptional Activation of miR-34a Contributes to p53-Mediated Apoptosis. Mol Cell. Cell Press; 2007;26:731–43.

110. Zhang L, Liao Y, Tang L. MicroRNA-34 family: A potential tumor suppressor and therapeutic candidate in cancer. J. Exp. Clin. Cancer Res. BioMed Central Ltd.; 2019. page 1–13.

111. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37:391–400.

112. Hahn M, Dambacher S, Dulev S, Kuznetsova AY, Eck S, Wörz S, et al. Suv4-20h2 mediates chromatin compaction and is important for cohesion recruitment to heterochromatin. Genes Dev. 2013;27:859–72.

113. Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, et al. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. Clin Cancer Res. 2008;14:7237–45.

114. Yokoyama Y, Matsumoto A, Hieda M, Shinchi Y, Ogihara E, Hamada M, et al. Loss of histone H4K20 trimethylation predicts poor prognosis in breast cancer and is associated with invasive activity. Breast Cancer Res [Internet]. 2014 [cited 2020 Jan 9];16:R66. Available from: http://breast-cancer-research.biomedcentral.com/articles/10.1186/bcr3681

115. Shinchi Y, Hieda M, Nishioka Y, Matsumoto A, Yokoyama Y, Kimura H, et al. SUV420H2 suppresses breast cancer cell invasion through down regulation of the SH2 domain-containing focal adhesion protein tensin-3. Exp Cell Res [Internet]. Elsevier; 2015;334:90–9. Available from: http://dx.doi.org/10.1016/j.yexcr.2015.03.010

116. Jørgensen S, Schotta G, Sørensen CS. Histone H4 Lysine 20 methylation: Key player in epigenetic regulation of genomic integrity. Nucleic Acids Res. 2013;41:2797–806.

117. Robert MF, Morin S, Beaulieu N, Gauthier F, Chute IC, Barsalou A, et al. DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. Nat Genet. Nature Publishing Group; 2003;33:61–5.

118. Suzuki M, Sunaga N, Shames DS, Toyooka S, Gazdar AF, Minna JD. RNA Interference-Mediated Knockdown of DNA Methyltransferase 1 Leads to Promoter Demethylation and Gene Re-Expression in Human Lung and Breast Cancer Cells. Cancer Res. American Association for Cancer Research; 2004;64:3137–43.

119. Kim H, Choi SY, Lim J, Lindroth AM, Park YJ. EHMT2 inhibition induces cell death in human non-small cell lung cancer by altering the cholesterol biosynthesis pathway. Int J Mol Sci. MDPI AG; 2020;21.

120. Wang L, Dong X, Ren Y, Luo J, Liu P, Su D, et al. Targeting EHMT2 reverses EGFR-TKI resistance in NSCLC by epigenetically regulating the PTEN/AKT signaling pathway. Cell Death Dis [Internet]. Nature Publishing Group; 2018 [cited 2020 Mar 16];9:129. Available from: http://www.nature.com/articles/s41419-017-0120-6

121. Tellez CS, Juri DE, Do K, Bernauer AM, Thomas CL, Damiani LA, et al. EMT and stem cell-like properties associated with miR-205 and miR-200 epigenetic silencing are early manifestations during carcinogen-induced transformation of human lung epithelial cells. Cancer Res. American Association for Cancer Research; 2011;71:3087–97.

122. Gan L, Yang Y, Li Q, Feng Y, Liu T, Guo W. Epigenetic regulation of cancer progression by EZH2: From biological insights to therapeutic potential [Internet]. Biomark. Res. BioMed Central Ltd.; 2018 [cited 2020 Mar 16]. page 10. Available from: https://biomarkerres.biomedcentral.com/articles/10.1186/s40364-018-0122-2

123. Zhang W, Lin J, Wang P, Sun J. miR-17-5p down-regulation contributes to erlotinib resistance in non-small cell lung cancer cells. J Drug Target [Internet]. Taylor and Francis Ltd; 2017 [cited 2020 Mar 15];25:125–31. Available from: https://www.tandfonline.com/doi/full/10.1080/1061186X.2016.1207647

124. Camidge DR, Pao W, Sequist L V. Acquired resistance to TKIs in solid tumours: Learning from lung cancer. Nat. Rev. Clin. Oncol. Nature Publishing Group; 2014. page 473–81.

125. Liao BC, Griesing S, Yang JCH. Second-line treatment of EGFR T790M-negative non-small cell lung cancer patients. Ther. Adv. Med. Oncol. SAGE Publications Inc.; 2019. page 1–16.

CHAPTER 2. LOSS OF SUV420H2 MEDIATES ERLOTINIB RESISTANCE IN NSCLC

Chapter overview

This chapter describes a genome-wide screen conducted to identify mediators of erlotinib resistance in NSCLC using the CRISPR-Cas9 system. Results of the screen led to the identification of a novel gene, SUV420H2, loss of which leads to the development of erlotinib resistance in NSCLC. Following which, two independent mechanisms by which loss of SUV420H2 causes erlotinib resistance have been delineated and validated.

2.1 Introduction:

Lung cancer is the leading cause of cancer-related mortality, accounting for 1.8 million deaths worldwide in 2018 (1). The majority of lung cancer patients are diagnosed with non-small cell lung cancer (NSCLC), a subtype that represents 85% of lung cancer cases. Since most lung cancer patients are diagnosed at later stages with metastatic disease surgical resection is not curative, and thus, the most effective treatment strategies are radiotherapy, chemotherapy, and targeted therapy. Targeted therapeutics are used based on the presence of particular molecular drivers, genes that the cancer cells are essentially addicted to. A few such drivers that are commonly present in NSCLC include KRAS, MEK, MET, HER2, and EGFR, many of which are either mutated or amplified in NSCLC, resulting in constitutive pro-growth signaling (2–7).

Epidermal Growth Factor Receptor (EGFR) is a cell surface receptor required for normal cell growth and proliferation. In 10-35% of NSCLC cases EGFR and its downstream pro-growth signaling pathways are constitutively activated due to mutations in the receptor, the most common of which are an amino acid substitution in exon 21 (L858R) or an in-frame deletion in exon 19. Mutant EGFR can be clinically targeted with a variety of Tyrosine Kinase Inhibitors (TKIs) including erlotinib and gefitinib, both of which are first generation TKIs, afatinib, a second generation inhibitor, or osimertinib a recently approved third generation TKI. In addition to inhibiting the activating EGFR mutations, osimertinib also specifically targets a secondary mutation in EGFR, T790M. Currently, EGFR-TKIs are front-line therapies for patients with tumors that have the respective EGFR mutations, but erlotinib has also been approved for use in

patients without EGFR mutations that have failed at least one prior chemotherapy regimen(8–11). Erlotinib binds reversibly and specifically to the ATP-binding pocket of EGFR with high efficacy, abrogating downstream growth and survival signaling pathways. While initially beneficial, within a year post treatment, patients develop resistance to erlotinib treatment, which is currently the major drawback of using such targeted therapies (12,13). Indeed, a similar response was observed with imatinib, another targeted inhibitor used to treat patients with chronic lymphocytic leukemia (14). The driver gene either incurs additional mutations or activates alternative signaling pathways to evade therapy. In the case of erlotinib treated patients that develop resistance, over 60% of patients suffer with a secondary mutation, T790M, whereas approximately 20% of patients undergo activation of bypass tracks that allow the tumor to evade targeting of the EGFR pathway through the use of alternative mechanisms that sustain their survival. Bypass tracks include signaling through oncogenic proteins such as MET, BRAF, HER2, PIK3CA, transformation of NSCLC into small cell lung cancer (SCLC), or through epithelial to mesenchymal transition (EMT) (15–19). In addition to a lack of understanding as to how the cell alters its growth dependency to these bypass tracks there are also approximately 15-20% of NSCLC tumors that acquire erlotinib resistance by mechanisms that remain unidentified (12,13).

While gain-of-function mechanisms that drive resistance have been identified, loss of tumor suppressive genes, such as PTEN, TP53, TET1, NF1 has also been reported to contribute to resistance (20–23). Indeed, many tumor suppressive proteins function as gatekeepers of the genome preventing spurious activation of oncogenes. To better define the genes involved in inhibiting the development of resistance, a genome-wide loss of function screen was conducted using the CRISPR-Cas9 system. Our data suggest that a novel epigenetic factor and *bona fide* tumor suppressor, never been reported to be involved in erlotinib resistance, SUV420H2 can be included among the gatekeepers of the genome. SUV420H2 can be included among the gatekeepers of the genome. SUV420H2 catalyzes the trimethylation of histone H4 at lysine-20 (H4K20) by utilizing mono-methylated H4K20 as a substrate, which is required for the establishment of heterochromatin and repression of genes (24–26). In addition to monomethylated H4K20 as a substrate, a few key players facilitate the sequential induction of H4K20 trimethylation modification of the genome; SUV39H2, another histone methyltransferase catalyzes H3K9me3 modification, recruiting the protein HP1 that physically associates with SUV420H2 to further mediate H4K20me3 of its substrate (24,27). Loss of SUV420H2 has previously been implicated

in causation of multiple cancers (28,29), but for the first time in this study we report that loss of SUV420H2 unravels two simultaneous mechanisms by which erlotinib resistance develops in NSCLC cells: 1) through de-repressing a long non-coding RNA, LINC01510 that ultimately upregulates the oncogene MET, and 2) through dampening DNA damage repair (DDR) mechanisms resulting in genomic instability.

2.2 Methods:

2.2.1 Cell culture:

All cell lines used in the study were obtained from American Type Culture Collection (ATCC), cultured and were confirmed to be free of mycoplasma. Cell lines generated during the study were authenticated by ATCC Cell Line Authentication. All cell lines were grown in RPMI media supplemented with 10% FBS and 1% Penicillin/Streptomycin. ECas9 cells were continuously cultured in media containing 1µg/ml Blasticidin, SUV420H2 mutant clones were grown in media containing 100ng/ml Puromycin, inducible-SUV420H2 Calu6 clones were cultured in 500ng/ml Puromycin containing media, and rescue clones were grown in media containing 100ng/ml Puromycin and 300µg/ml G418 containing media.

2.2.2 Drug Preparation for in vitro studies:

Erlotinib (S7786, Selleck Chemicals), afatinib (850140-72-6, Sigma Aldrich), gefitinib (S1025, Selleck Chemicals), and osimertinib (S7297, Selleck Chemicals) were dissolved in DMSO to prepare 0.4 M stock solutions, which were aliquoted and stored in -80°C. Cisplatin was dissolved in water (1mg/ml) and stored at room temperature. A 200 µM working dilution of all the drugs was prepared in complete medium and were used to prepare the indicated concentrations for all *in vitro* experiments.

2.2.3 Knock-out CRISPR screen:

EKVX cells ($4X10^5$) were plated in 6-well plates and were transfected with 3µg of linearized lentiCas9-Blast (Addgene, 52962) using lipofectamine 2000 (11-668-019, Thermo Fisher Scientific), as per manufacturer's instructions. Forty-eight hours later, cells were selected using

5µg/ml Blasticidin. ECas9 (clone 7) cells stably expressing Cas9 plasmid were clonally selected and characterized. Lentiviral sgRNA library (A and B) were generated and the titer was determined as previously described (30). The GeCKO V2 library has 6 sgRNAs targeting protein coding gene and 4 sgRNAs targeting miRNAs. To achieve a 300-fold coverage of the libraries seventeen 12-well plates were each seeded with $4.5X10^5$ ECas9 cells. Nine plates were transduced with library A, and 8 plates were transduced with library B, both at a multiplicity of infection (MOI) of 0.4 in the presence of Polybrene (10µg/ml). Twenty-four hours post transduction, cells were pooled and ~1.31×10^7 cells were re-plated in each of seven 15 cm plates containing complete media supplemented with 2µg/ml Blasticidin. Forty-eight hours later cells were plated in six 15 cm plates in media containing 2µg/ml Puromycin, to select for library-transduced cells, and 2µg/ml Blasticidin. Seventy-two hours later, 2.6×10^7 cells were stored for baseline and 2.6×10^7 cells were re-plated. The following day, media was replaced with GI75 erlotinib containing media (1.23µM erlotinib) and cells were continuously exposed to GI75 erlotinib for 15 passages. Three biological replicates were performed, and genomic DNA from each baseline and erlotinib treated sample were isolated using the Genomic DNA isolation kit (K1820-01, Thermo Fisher Scientific) following the manufacturer's protocol. For sequencing library preparation, two sequential PCR reactions were conducted for each sample. The first PCR reaction specifically amplified sgRNAs from 1µg of gDNA isolated from each sample. Twenty-five such PCR reactions were conducted, pooled, and gel purified using QIAEX II Gel Extraction Kit (20021, Qiagen). Each PCR1 reaction product (10 ng) was then used for each of 20 PCR2 reactions that were pooled and gel purified. PCR2 fragment sizes and library quality were evaluated on a bioanalyzer (Agilent). Both PCR1 and PCR2 primers are listed in **Table 2.1** (Integrated DNA Technologies). Barcodes included in PCR2 primers were used to identify the samples after deep sequencing. All sequencing was conducted using a NovaSeq 6000 (Illumina). FastQC version 0.11.7 was used to observe sequencing data quality before and after trimming. Cutadapt version 1.13 was used to trim adapters from reads. Reads post-trimming that were shorter than 18nt were discarded. MAGeCK-VISPR v. 0.5.6 was used to perform mapping, allowing no mismatches to ensure accuracy and to reduce bias. Finally, MAGeCK was used to identify over- and under-represented sgRNAs in treated samples relative to baseline, represented as β-scores(31,32).

Table 2.1.Primer sequences utilized to conduct the CRISPR-Cas9 knock-out screen. Multiple PCR2 primers were used, each with an independent barcode that allows for sorting of sample specific sgRNAs post sequencing.

PCR	Sample	Primer name	Primer direction	Primer sequence
PCR 1	All samples	1st PCR primer	Forward	TCTTTCCCTACACGACGCTCTTCCGATCTNNNNAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG
		1st PCR primer	Reverse	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCC
PCR2	EKVX-Baseline 1	UDA5050	Forward	AATGATACGGCGACCACCGAGATCTACACTGACAATGTCACACTCTTTCCCTACACGAC
		UDA7143	Reverse	CAAGCAGAAGACGGCATACGAGATAGAAGCCAATGTGACTGGAGTTCAGACGTG
	EKVX-Replicate 1	UDA5051	Forward	AATGATACGGCGACCACCGAGATCTACACCGACCTAACGACACTCTTTCCCTACACGAC
		UDA7142	Reverse	CAAGCAGAAGACGGCATACGAGATGACTCACTAAGTGACTGGAGTTCAGACGTG
	EKVX-Baseline 2	UDA5052	Forward	AATGATACGGCGACCACCGAGATCTACACTAGTTCGGTAACACTCTTTCCCTACACGAC
		UDA7141	Reverse	CAAGCAGAAGACGGCATACGAGATAGTCTGTCGGGTGACTGGAGTTCAGACGTG
	EKVX-Replicate 2	UDA5053	Forward	AATGATACGGCGACCACCGAGATCTACACGCCGCACTCTACACTCTTTCCCTACACGAC
		UDA7140	Reverse	CAAGCAGAAGACGGCATACGAGATGTATTCTCTAGTGACTGGAGTTCAGACGTG
	EKVX-Baseline 3	UDA5054	Forward	AATGATACGGCGACCACCGAGATCTACACATTATGTCTCACACTCTTTCCCTACACGAC
		UDA7139	Reverse	CAAGCAGAAGACGGCATACGAGATACGCCTCTCGGTGACTGGAGTTCAGACGTG
	EKVX-Replicate 3	UDA5055	Forward	AATGATACGGCGACCACCGAGATCTACACAGAACCGAGTACACTCTTTCCCTACACGAC
		UDA7138	Reverse	CAAGCAGAAGACGGCATACGAGATTAACCGCCGAGTGACTGGAGTTCAGACGTG

2.2.4 Knockout, knockdown, overexpression and rescue experiments:

To generate the SUV420H2 sgRNA, two oligos (see **Table 2.2**) were annealed and 5' phosphorylated (T4 Polynucleotide Kinase kit, M0201S, NEB) as described previously (33). Simultaneously, the CRISPR-Cas9 plasmid, LentiCRISPRv2 (Addgene, 52961) was digested using BsmBI (R0580, NEB), dephosporylated (Antarctic phosphatase, M0289S, NEB) and gel purified using QIAEX II Gel Extraction Kit (20021, Qiagen). The annealed oligos were ligated into the gel purified vector, transformed into Stabl3 bacteria and miniprepped, as outlined previously (33). Three micrograms of the so generated pLV-sgSUV420H2 plasmid was linearized and forward transfected in $4X10^5$ ECas9 cells using lipofectamine 3000 (L3000015, Thermo Fisher Scientific), following the manufacturer's protocol to generate SUV420H2 mutant clones A, C, E.

For all siRNA-mediated knockdown experiments, 30nM of the respective siRNAs were reverse transfected into 10,000 (for dose curves and proliferation assays) or in $4X10^5$ SUV420H2 mutant clones using Lipofectamine RNAiMAX (13-778-150, Thermo Fisher Scientific) following the manufacturer's protocol. siRNAs used in the study: siMET (Catalog # 4390824, Assay ID # s8700, Thermo Fisher Scientific) and siLINC01510 (Catalog #: 4392420, Assay ID # n506737 Thermo Fisher Scientific).

For generation of DOX inducible overexpression plasmid, the SUV420H2 sequence was amplified from an ORF expression clone for SUV420H2 (eGFP tagged) (EX-V0810-M98, GeneCopoeia) introducing a a stop codon. The sequence was purified and ligated into the pLVX-Tetone. The oligonucleotides used to perform the sequence exchange are indicated in **Table 2.2**. Following construction of the pLVX-Tetone-SUV420H2 plasmid, 3µg of the linearized plasmid was transfected into $4X10^5$ Calu6 cells using lipofectamine 3000 to generate the SUV420H2-inducible Calu6 clone.

Next, to generate the rescue lines from SUV420H2 mutant clone C, a puromycin resistance gene was cloned into pLVX-Tetone-SUV420H2 using the primers outlined in **Table 2.2**. Following generation of the pLVX-Tetone-SUV420H2-puro plasmid, 3µg of the linearized plasmid was transfected in $4X10^5$ SUV420H2 mutant cells using lipofectamine 3000 for the generation of inducible-SUV420H2 rescue clones R1, and R2.

Finally, to test effect of MET or LINC01510 on erlotinib resistance, pT3-EF1a-c-Met (31784, Addgene) or pCMV-Hygro-LINC01510 (Twist Bioscience) were transfected using Lipofectamine 3000 in $4X10^5$ ECas9 cells.

2.2.5 Genotyping of mutation:

Genomic DNA was amplified in the region containing the expected *KMT5C* mutation using Q5 high fidelity polymerase (M0491L, NEB). Primers for amplification and sequencing are outlined in **Table 2.2**.

2.2.6 Bioinformatic analysis of TCGA data:

Gene Expression Profiling Interactive Analysis (GEPIA) database (34) (*http://gepia.cancer-pku.cn/*) was used to evaluate *KMT5C* and *LINC01510* levels in NSCLC patient samples and non-tumorigenic controls. GEPIA was also used to perform a correlation analysis between *LINC01510* and *MET* levels. TANRIC (35) (*http://bioinformatics.mdanderson.org/public-software/tanric/*) was used to determine the prognostic value of *LINC01510* in NSCLC. Integrated Genome Viewer (IGV 2.3) was used to view bed files reported by GSE59316 using Human genome 19 (hg19) browser.

2.2.7 Western Blot:

Four-hundred thousand cells were grown in individual wells of a 6 well plate, and lysates were isolated at time points specified in figure legends using RIPA buffer (Sodium chloride (150 mM), Tris-HCl (pH 8.0, 50mM), N P-40 (1 %), Sodium deoxycholate (0.5 %), SDS (0.1 %), ddH2O (up to 100 mL)) containing 1X protease inhibitor cocktail (PIA32955, Thermo Fisher Scientific). Protein quantification was performed using Pierce BCA Protein Assay kit. Equal amounts of protein lysate were resolved through 12% or 4-20% polyacrylamide gels and transferred onto a polyvinylidene difluoride (PVDF) membrane. Membranes were blocked using LI-COR buffer for 1 hour at room temperature, and incubated overnight in primary antibody at 4°C. The primary antibody was detected using 1:800 IR 800CW secondary antibody. Blots were scanned, and data quantified using the Odyssey LI-COR imaging system and software. Antibodies used: mouse

H4K20me3 (39672; Active Motif), rabbit H4K20me3 (ab9053, abcam), rabbit MET (D1C2) XP (8198, Cell Signaling), mouse β-ACTIN (3700, Cell Signaling)

2.2.8 In-Cell Western:

Ten-thousand cells were grown in individual wells of a 96-well plate. Forty-eight hours post plating, cells were fixed using cold 100% methanol for 20 minutes at 4 C. Post fixing, cells were permeabilized using 0.2% TritonX in 1X PBS at room temperature for 30 minutes. Cells were blocked using LI-COR blocking buffer for 1.5 hours followed by overnight incubation with primary antibody at 4°C. The primary antibody was detected using 1:800 IR 800CW secondary antibody (LI-COR). The IR-800 signal was quantified using the Odyssey LI-COR imaging system and software. Antibodies used: 1:400 mouse H4K20me3 (39672, Active Motif), 1:500 rabbit GAPDH (2118, Cell Signaling)

2.2.9 Immunofluorescence:

Two-hundred thousand cells were seeded on collagen coated (overnight with 150ng/µl collagen) coverslips that were arranged in individual wells of a 12-well plate. Forty-eight hours post plating, cells were fixed using cold 100% methanol for 20 minutes in 4°C. Post fixing, cells were permeabilized using 0.2% TritonX in 1X PBS at room temperature for fifteen minutes. Following which, cells were blocked using LI-COR blocking buffer for 1 hour followed by overnight incubation with 1:50 mouse H4K20me3 (39672, Active Motif) or anti-mouse 1:400 γ-H2AX (05636, Millipore) at 4°C. 1:500 anti-mouse Alexa Fluor 647 (A-31571, Thermo Fisher Scientific) was used to detect H4K20me3 and 1:1000 Hoechst dye (H3570, Thermo Fisher Scientific) for 2 hours at room temperature. Coverslips were mounted on glass slides using ProLong Glass Antifade Mountant (P36982, Thermo Fisher Scientific). Fluorescence images were collected using the Nikon A1RMP microscope and analyzed using NIS-Elements Microscope Imaging Software.

2.2.10 RNA isolation and Quantitative real time PCR (qRT-PCR):

Four-hundred thousand cells were grown in individual wells of a 6-well plate, and total RNA was isolated after 48 or 96 hours, as indicated, using the miRneasy Kit (217004, Qiagen) according to the manufacturer's instruction. DNase I digestion (79254, Qiagen) was used in each RNA

purification reaction to remove genomic DNA. RNA integrity was evaluated on a 1.5% agarose gel, and total RNA quantified using a nanodrop. cDNA was then synthesized from 1μg of total RNA using MiScript Reverse Transcriptase kit (218161, Qiagen), as indicated by the manufacturer's protocol. Q-RT-PCR was conducted using the miScript SYBR Green PCR Kit (218073, Qiagen) as indicated by the manufacturer's protocol, to quantify target gene mRNA expression. The following primers were obtained: *GAPDH* (loading control) (QT00079247, Qiagen), *LINC01510* (LPH09040A, Qiagen), and MET (QT00023408, Qiagen). Primers for *KMT5C* quantification are indicated in **Table 2.2**.

2.2.11 Chromatin Immunoprecipitation - quantitative PCR (ChIP-qPCR):

Briefly, a total of $2X10^7$ cells were fixed using 1% of filter-sterilized formadehyde for 10 minutes at room temperature. The formaldehyde was quenched with 2.5M Glycine (55μL per ml of media) for 5 min. Cells were washed with cold PBS and scraped into fresh cold PBS. Cells were pelleted by centrifuging at 1500 rpm for 10 minutes at 4°C. The cell pellet was resuspended in 10 mL of freshly prepared cold cell lysis buffer (5mM PIPES, 85mM KCl, 0.5% NP40), kept on ice for 10 minutes followed by centrifuging at 1000 rm for 10 minutes at 4°C. The lysed cells were resuspended in 1 mL of nuclei lysis buffer (50mM Tris-HCl (pH 8.0),10mM EDTA, 1% SDS) containing 0.1% protease inhibitor cocktail (PIA32955, Thermo Fisher Scientific) and were transferred into 2mL eppendorf tubes, on ice. Cross-linked chromatin from the isolated nuclei was sonicated using a probe sonicator (60% duty cycle) for 10 seconds with a 1 minute rest, for 15 cycles to fragment DNA (100-500 bps). Fragmented DNA was immunoprecipitated with antibodies against mouse H4K20me3 (39672, Active Motif), or negative control mouse IgG (5415, Cell Signaling Technology) at 4°C overnight with gentle rotation. The immunoprecipitated DNA was purified using the DNA isolation kit (K1820-01; Thermo Fisher Scientific) following manufacturer's protocol. DNA was used as a template for qRT-PCR as described above. All primer sequences used for qRT-PCR are listed in **Table 2.2**. ChIP data are presented as fold enrichment of DNA immunoprecipitated with H4K20me3 relative to values obtained for DNA immunoprecipitated with IgG control.

Table 2.2. Primers utilized in the study. Designed and purchased from Integrated DNA Technologies.

Primer use		**Primer direction**	**Primer sequence**
pLV-sgSUV420H2		Forward	CACCGCGGCCCGCTACTTCCAGAGC
		Reverse	AAACGCTCTGGAAGTAGCGGGCCGC
pLVX-Tetone-SUV420H2		Forward	TCGTAAAGAATTCACCATGGGGCCCGACA GAGTGACAGCA
		Reverse	GAGATCTGGATCCTCAGTACAGCTCTTCA CCGCCGAC
pLVX-Tetone-SUV420H2-puro		Forward	CCGCTACGCGTTCAGAAGAACT
		Reverse	AGCGGCGTACGATGATTGAACA
KMT5C genomic locus amplification		Forward	GAGCAGATGGGAGGTGCGGCGACAGT
		Reverse	GAGCTCAGAAGAAAGGAGACAGAT
KMT5C genomic locus sequencing		Forward	CCTCTCCTTAGCCTGGTCCT
		Reverse	CAAGGGCTAGGAAGTCAGGG
KMT5C quantification		Forward	TCGGTTTCCGCACCCATAAG
		Reverse	CGGAGGTAGCGATAGACGTG
ChIP-QPCR	FOXA1 mark	Forward	AAGGAGAGGTGCGTTGTTTG
		Reverse	CATTCTCCCACGAAAGGCAG
	FOXA1 exon	Forward	AAGACTCCAGCCTCCTCAAC
		Reverse	CGGGTGGTTGAAGGAGTAGT
	Linc01510 mark	Forward	GCTTCTTGTCCCTCCCAGAT
		Reverse	GCAGAAGTGAGAGGAAGGGT
	Up 1	Forward	CACACTGGAGTTCTTGCCAC
		Reverse	TATGCACTCCTTCACTGGGG
	Up 2	Forward	GCAGTCCAGCTAAGCAATCC
		Reverse	GACATCTTGGGAAGGGGACA
	Up 3	Forward	CCTCTTCACATCCCACAGGT
		Reverse	CTCTGCTGGCTTGATCATTG
	MET	Forward	GATCAAGGAAATGGGGCGTT
		Reverse	GGGACTAGGGCCTATTGTCA
	Down 1	Forward	CCCTGCCTCTCATCAACTGA
		Reverse	GTTGAGCCACTAAACCACCC
	Down 2	Forward	TGCCTGGTCTCCTGTTAACA
		Reverse	ATCTGTCTTCTCCCTGTGCC
	Down 3	Forward	AGTCCAAGATCAAGGCACCA
		Reverse	AGGCCTTTCTTGTACCCCTT

2.2.12 Erlotinib dose response:

The protocol followed to evaluate erlotinib dose response was as per the NCI-60 Cell Five-Dose Screen (NCI-60, DTP (36)). Briefly, Sulforhodamine B colorimetric assay (SRB assay)(37) was performed by exposing cells to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. To normalize data, percent of cells was calculated based on first correcting for the number of cells at the start of the assay (time zero = tz), followed by normalization of cell number to respective corrected DMSO values.

2.2.13 Proliferation:

Ten thousand NSCLC cells or transfected cells were seeded in 6 replicates in wells of a 96-well plate, which was placed in a live-imaging system, Incucyte s3 2018A (ESSEN BioScience). Plates were incubated in the system for the specified times. Four images per well were obtained every 2 hours using the 10X objective. Confluence was evaluated using Incucyte s3 2018A software. To normalize data, percent of cells was calculated based on first correcting for the number of cells at the start of the assay (time zero = tz), followed by normalization of cell number to respective corrected DMSO values. Data is represented relative to controls, as described in figure legends.

2.2.14 Comet assay:

Glass slides preheated at 50°C and subsequently coated with 1.5% normal melting agarose, were placed in 4 °C overnight. The following day, cells were trypsinized and resuspended in PBS such that the concentration is10,000 cells per 10µL. The suspension of cells was thoroughly mixed with 75µL of low melting point agarose (LMPA) which was maintained at 37°C and added onto the overnight 1.5% normal melting agarose coated glass slide. A coverslip was placed on top of the agarose- cell amalgam on the glass slide, and placed on ice until the agarose solidified (approximately 30 minutes). Coverslips were removed and slides were incubated in lysis solution (2.5M NaCl, 100mM EDTA, 10mM Trizma Base) overnight at 4°C. Slides were placed in an electrophoresis tank containing freshly prepared electrophoresis alkaline buffer (300 mM NaOH, 1 mM EDTA) and were allowed to acclimate for 20 minutes prior to being resolved at 24V, for 30min at 4 °C. Neutralization buffer (0.4 M Tris, pH 7.5) was added dropwise to coat the slides followed by a 5min incubation. Excess buffer was blotted, and the neutralization was repeated 3

times. The slides were stained with gel red for 5min and excess stain was removed by blotting. Comets were scored immediately using Nikon Olympus IX73 fluorescent microscope.

2.2.15 Statistical analysis:

All data were analyzed using GraphPad Prism version 7 software (GraphPad Software) and are presented as mean values ± standard deviation (SD). Pearson's correlation was utilized to evaluate linear correlation between SUV420H2 and/or H4K20me3 and GI50 erlotinib values. Student's t-test or one-way ANOVA were performed, as specified in the figure legends. P-value of < 0.05 was considered significant.

2.3 Results:

2.3.1 Identification of novel mediators of erlotinib resistance

To identify genes, that when mutated confer resistance to erlotinib sensitive cells, a genome-wide CRISPR-Cas9 knock-out screen was performed. The screen was conducted in EKVX cells, a cell line that was determined to be erlotinib sensitive in data obtained from the Developmental Therapeutics Program, which is maintained by the National Cancer Institute (NCI-60, DTP (36)). The cells were engineered to stably express the Cas9 protein and resulting clones were validated for their response to erlotinib, which was similar to the parental EKVX cells **(Supplementary Figure 2.1).** A single Cas9-expressing EKVX clone was taken forward to conduct the screen, which is hereafter referred to as ECas9. The ECas9 cells were infected with the GeCKO V2 sgRNA lentiviral library targeting 19,052 protein-coding genes and 1,864 miRNA genes (**Figure 2.1A**)(38). To obtain full coverage of the lentiviral sgRNA library, transduction was performed at 300-fold coverage and was conducted in triplicates to mitigate false positives. Post transduction, sgRNA integration was determined in a subset of cells prior to selection to identify the baseline level for each sgRNAs. The remaining cells were grown in the presence of 75% growth inhibitory concentration of erlotinib (GI75) for 15 passages followed by sgRNA identification in the selected population. Subsequently, sgRNA representation in the GI75 treated cells and the baseline cells was evaluated by high-throughput sequencing. Combined analysis of the three replicates using the MAGeCK-VISPR algorithm identified significantly enriched sgRNAs in the population of cells that were cultured in erlotinib (**Table 2.3**, **Figure 2.1B**) (31,32). Following the analysis, multiple

genes that were previously reported to be 1) downregulated during acquired resistance to chemotherapy treatment (TKI or non-TKI) (39–44), 2) highly expressed in erlotinib sensitive cells (45,46), and 3) *bona fide* tumor suppressors (29,47–53), were among the top hits in the screen supporting the validity of the screen.

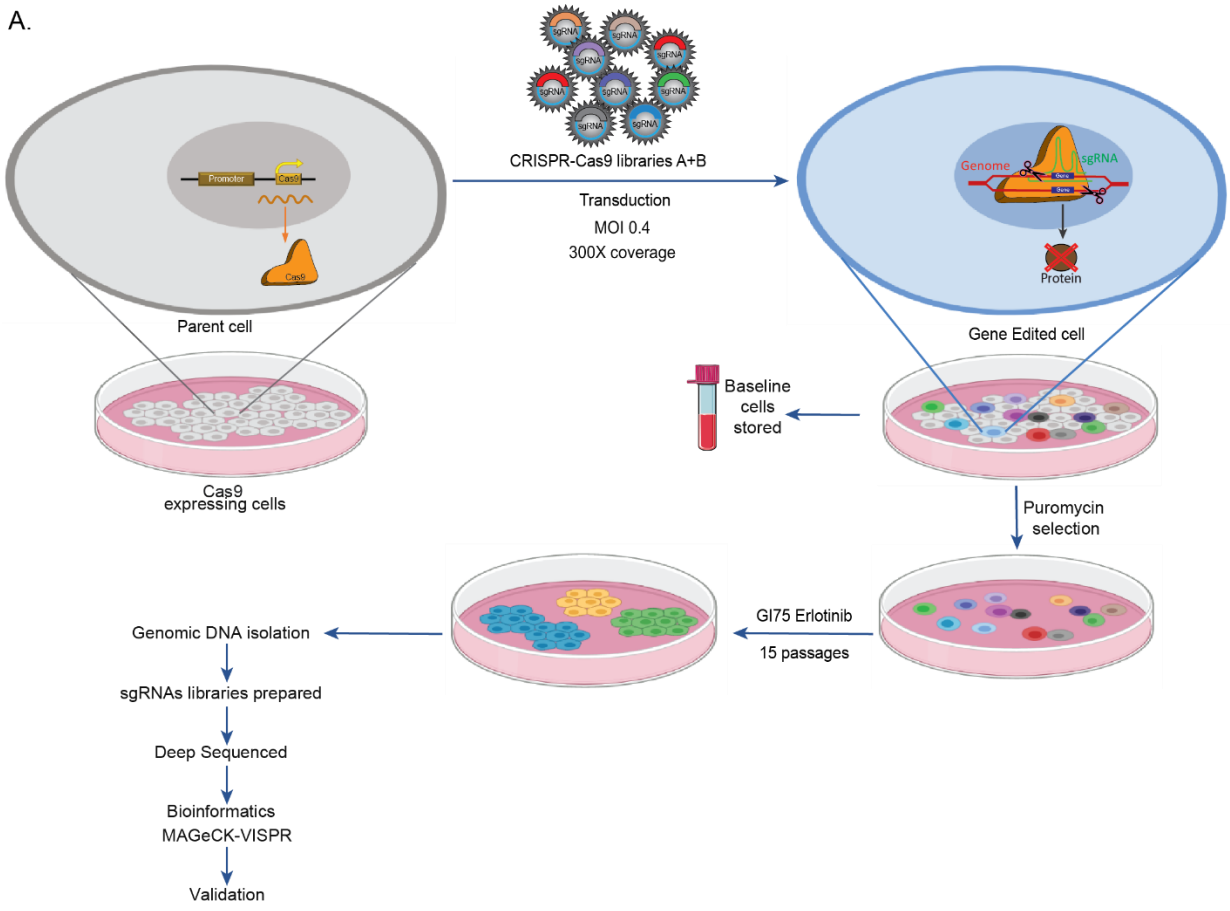
2.3.2 Low expression of SUV420H2 is associated with erlotinib resistance, and predicts poor prognosis in NSCLC

The top hit from the CRISPR-Cas9 knock-out screen SUV420H2, is a histone methyltransferase encoded by *KMT5C.* SUV420H2 specifically trimethylates histone H4 lysine-20 (H4K20), which is associated with transcriptional repression and is important for establishing constitutive heterochromatic regions (24,25,54,55). Multiple studies have reported on the role of SUV420H2 as a tumor suppressor, and both SUV420H2 and H4K20 trimethylation (H4K20me3) are severely downregulated in multiple cancers (28,29,48,56–58). To determine if SUV420H2 is also a mediator of erlotinib response, various validation assays were performed. Firstly, using a panel of NSCLC cell lines, a negative correlation between *KMT5C* and erlotinib response was determined (**Figure 2.2A-C,** Pearson r = -0.83). Due to the lack of a sensitive and specific SUV420H2 antibody, we evaluated the downstream effector of SUV420H2, H4K20me3 as a proxy for SUV420H2 activity (**Supplementary Figure 2.2A, B)**. Indeed, in the same cell line panel, we determined that H4K20me3 levels positively correlate with *KMT5C* (Pearson r = 0.6905, **Supplementary Figure 2.2C**). On the other hand, similar to the negative correlation between *KMT5C* and erlotinib response in the NSCLC panel, H4K20me3 also displayed a negative correlation with erlotinib response (Pearson r = -0.61, **Supplementary Figure 2.2D**). These strong correlations suggest a possible role for SUV420H2 in mediating the response of NSCLC cells to erlotinib.

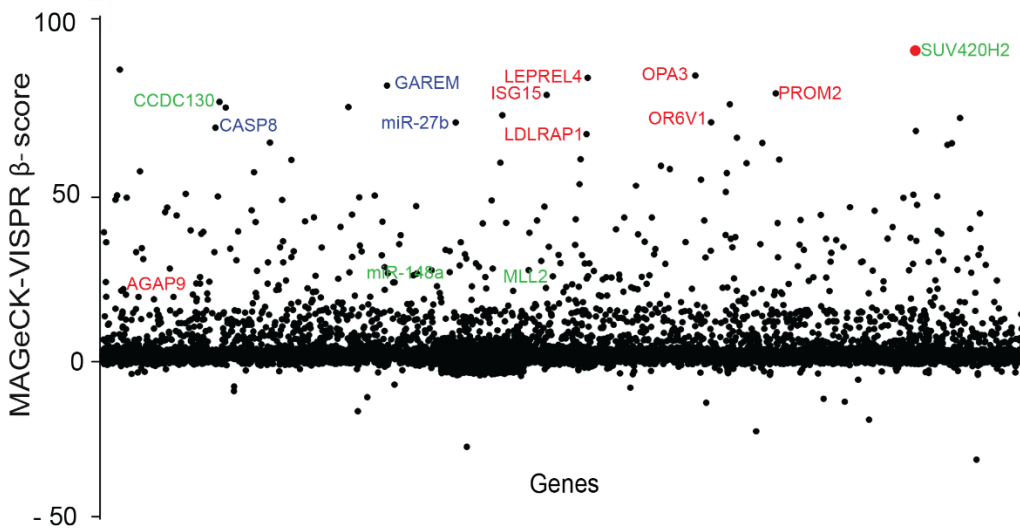
Next, we investigated *KMT5C* levels in NSCLC patient samples using publicly available data provided in The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) project (34,59,60). Patient samples were compared to non-cancerous control tissues using Gene Expression Profiling Interactive Analysis (GEPIA, **Figure 2.2D**). *KMT5C* levels trended were lower in both lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) samples relative to normal samples.

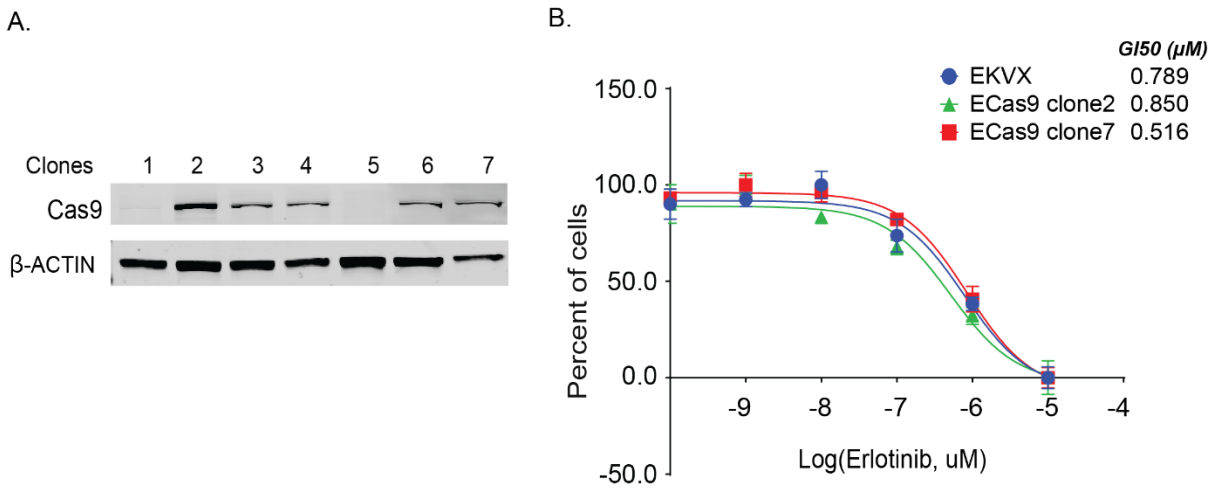
Figure 2.1.A genome-wide CRISPR-Cas9 screen identifies mediators of erlotinib resistance in NSCLC. A) Outline of the CRISPR-Cas9 knock-out screen. B) Fold enrichment (β-score) analysis of sgRNAs targeting genes present in the CRISPR-Cas9 lentiviral library, using MAGeCK-VISPR analysis. Genes represented in blue have previously been reported to be are downregulated in cells post TKI treatment, those in red were reported to be high in erlotinib sensitive cells (CtRPv2), and those depicted in green are published tumor suppressors.

A.


sgRNA
Promotor
Cas9
Parent cell
CRISPR-Cas9 libraries A+B
Transduction
MOI 0.4
300X coverage
Genome
sgRNA
Protein
Gene Edited cell
Baseline
cells
stored
Cas9
expressing cells
Puromycin
selection
GI75 Erlotinib
15 passages
Genomic DNA isolation
sgRNAs libraries prepared
Deep Sequenced
Bioinformatics
MAGeCK-VISPR
Validation

B.


100
50
0
- 50
MAGeCK-VISPR β- score
Genes
SUV420H2
CCDC130
CASP8
GAREM
miR-27b
LEPREL4
ISG15
LDLRAP1
OPA3
OR6V1
PROM2
AGAP9
miR-142a
MLL2



Supplementary Figure 2.1. Characterization of single Cas9 expressing EKVX clones. A) Western Blot analysis of Cas9 levels in EKVX cells stably expressing Cas9. β-ACTIN was used as a loading control. B) Parental EKVX cells or the ECas9 clones 2 and 7 were exposed to varying concentrations of erlotinib or the highest equivalent volume of dimethyl sulfoxide (DMSO, negative control) containing media for 72 hours. Erlotinib dose response was evaluated using the SRB assay. Post-normalization, GI50 concentration of erlotinib was calculated from their respective dose curves for each cell line.

Table 2.3. Candidate genes identified from the CRISPR-Cas9 knock out screen. The thirty-five hits identified by MAGeCK-VISPR analysis and β-score, p-value, and false discovery rate (FDR) are enlisted.

Target	β-score	p-value	FDR
SUV420H2	97	8.30E-05	0.07
ADSS	91	0.00021	0.07
OPA3	89	0.00028	0.07
LEPREL4	88	0.00032	0.07
GAREM	86	0.00049	0.07
ISG15	83	0.00065	0.07
PROM2	83	0.00065	0.07
hsa-mir-602	77	0.00082	0.07
CCDC130	81	0.00088	0.07
PCSK2	80	0.00091	0.07
FAM120AOS	79	0.001	0.07
CCL23	79	0.0011	0.07
TNFSF12	76	0.0028	0.07
hsa-mir-27b	74	0.0081	0.11
SMN2	25	0.012	0.16
OR6V1	74	0.012	0.16
SYBU	72	0.012	0.17
CASP8	73	0.012	0.17
LDLRAP1	71	0.013	0.17
PFDN2	70	0.013	0.17
CPA3	68	0.013	0.17
PP2D1	68	0.013	0.17
TMEM234	68	0.013	0.17
TMEM147	67	0.013	0.17
hsa-mir-5699	62	0.016	0.21
hsa-mir-512-1	50	0.016	0.21
KMT2D/ MLL2	22	0.016	0.21
hsa-mir-648	43	0.016	0.21
AGAP9	22	0.016	0.21
hsa-mir-4669	43	0.016	0.21
RPL41	38	0.016	0.21
hsa-mir-3183	37	0.016	0.21
hsa-mir-1268a	34	0.017	0.22
hsa-mir-147b	34	0.017	0.22
hsa-mir-148a	27	0.018	0.24

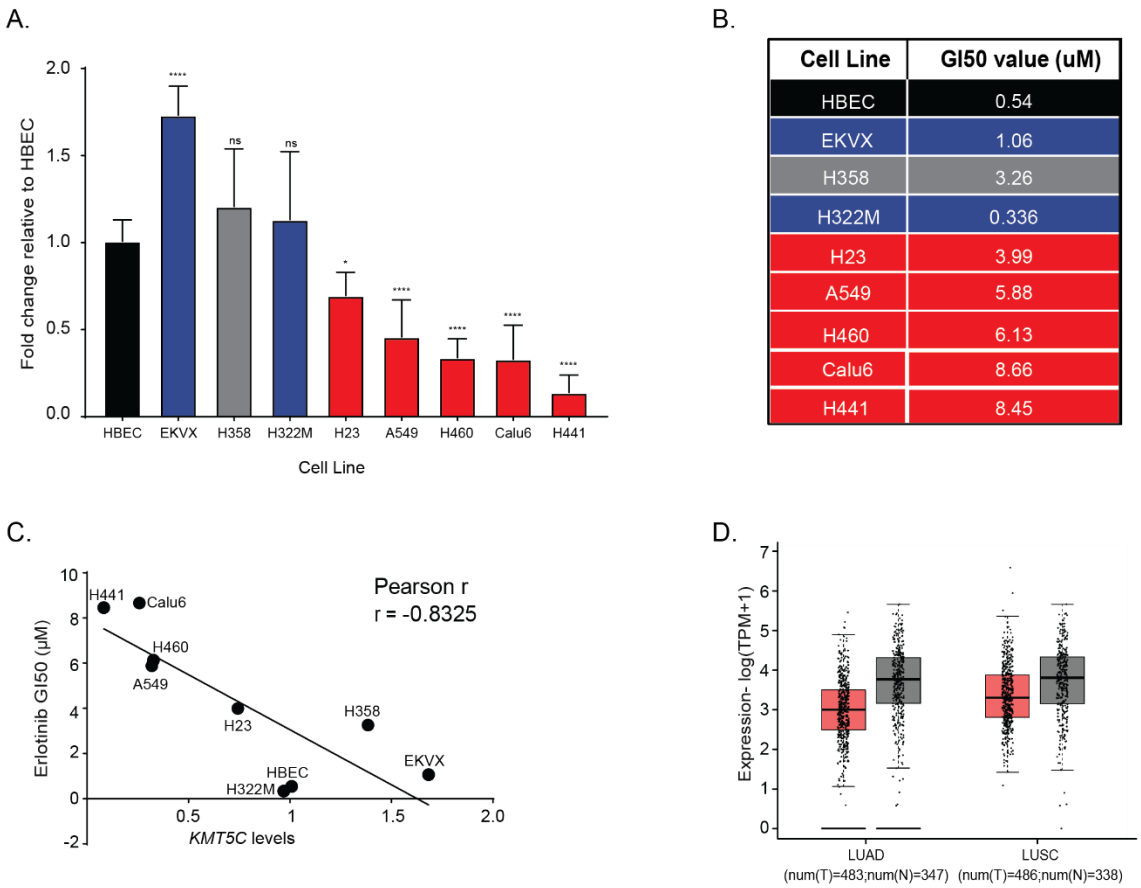
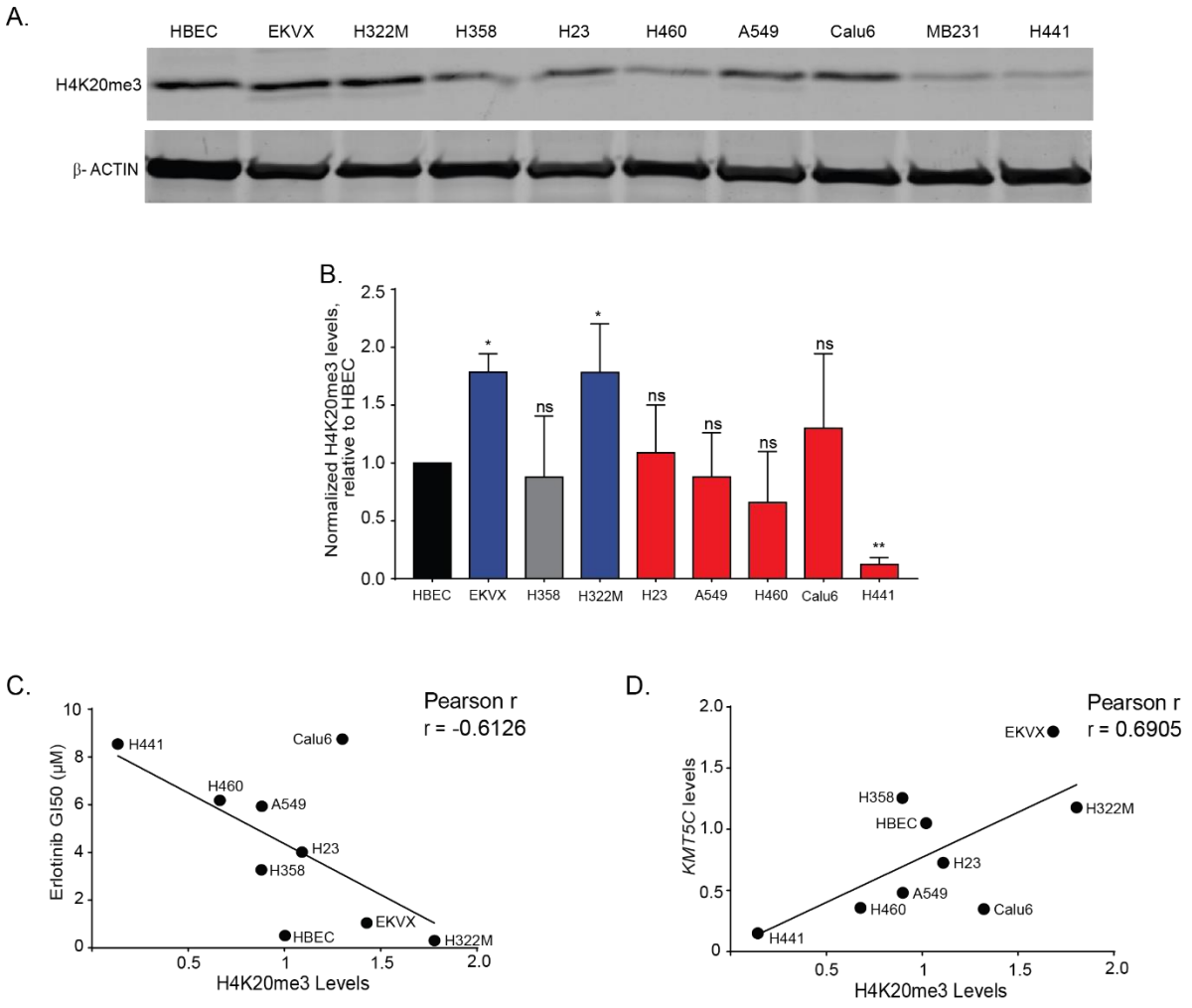


Figure 2.2. Reduced SUV420H2 expression correlates with erlotinib resistance in NSCLC cells, and poor prognosis in NSCLC patients. A) Endogenous expression of *KMT5C* in a panel of NSCLC cells, relative to a normal lung epithelial cell line (Human Bronchial Epithelial Cells, HBEC), evaluated by qRT-PCR. GAPDH is utilized as the endogenous control. One-way ANOVA followed by Dunnett's Multiple Comparison test was used to evaluate statistical significance of SUV420H2 transcript levels relative to HBEC cells. Color of bars represents previously reported response to erlotinib in the DTP study. Red bars indicated cells lines resistant to erlotinib, blue bars represent cell lines reported to be sensitive. Neither HBEC (non-tumorigenic control) or H358 (NSCLC) were included in the DTP dataset. B) erlotinib dose response via SRB assay was evaluated by exposing each of the individual cell lines to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. GI50 concentrations of erlotinib were calculated from each cell lines respective dose curve. Colors are as in A. C) Correlation analysis between *KMT5C* from A and GI50 erlotinib concentrations from B was evaluated using Pearson correlation test. D) GEPIA analysis for *KMT5C* levels in normal and tumor samples from LUAD and LUSC data obtained from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) databases. TPM= Transcripts per million, T= Tumor, N=Normal



Supplementary Figure 2.2. Reduced H4K20me3 correlates with erlotinib resistance in NSCLC cells. A) Representative western blot image of H4K20me3 levels in a panel of NSCLC cells. β-ACTIN was used as a loading control. MB231, a breast cancer cell line was included as a positive control line with endogenously low levels of SUV420H2 (37,50). B) H4K20me3 levels relative to β-ACTIN normalized to HBEC derived from four biological replicates. One-way ANOVA followed by Dunnett's Multiple Comparison test was used to evaluate statistical significance of H4K20me3 levels relative to HBEC cells. C) Correlation analysis between quantified H4K20me3 levels from panel B and GI50 erlotinib values from Figure 2B. D) Correlation analysis between quantified H4K20me3 levels from panel B and *KMT5C* from Figure 2A. Evaluations in C and D were conducted using the Pearson correlation test.

2.3.3 Loss of SUV420H2 confers resistance to EGFR inhibitors

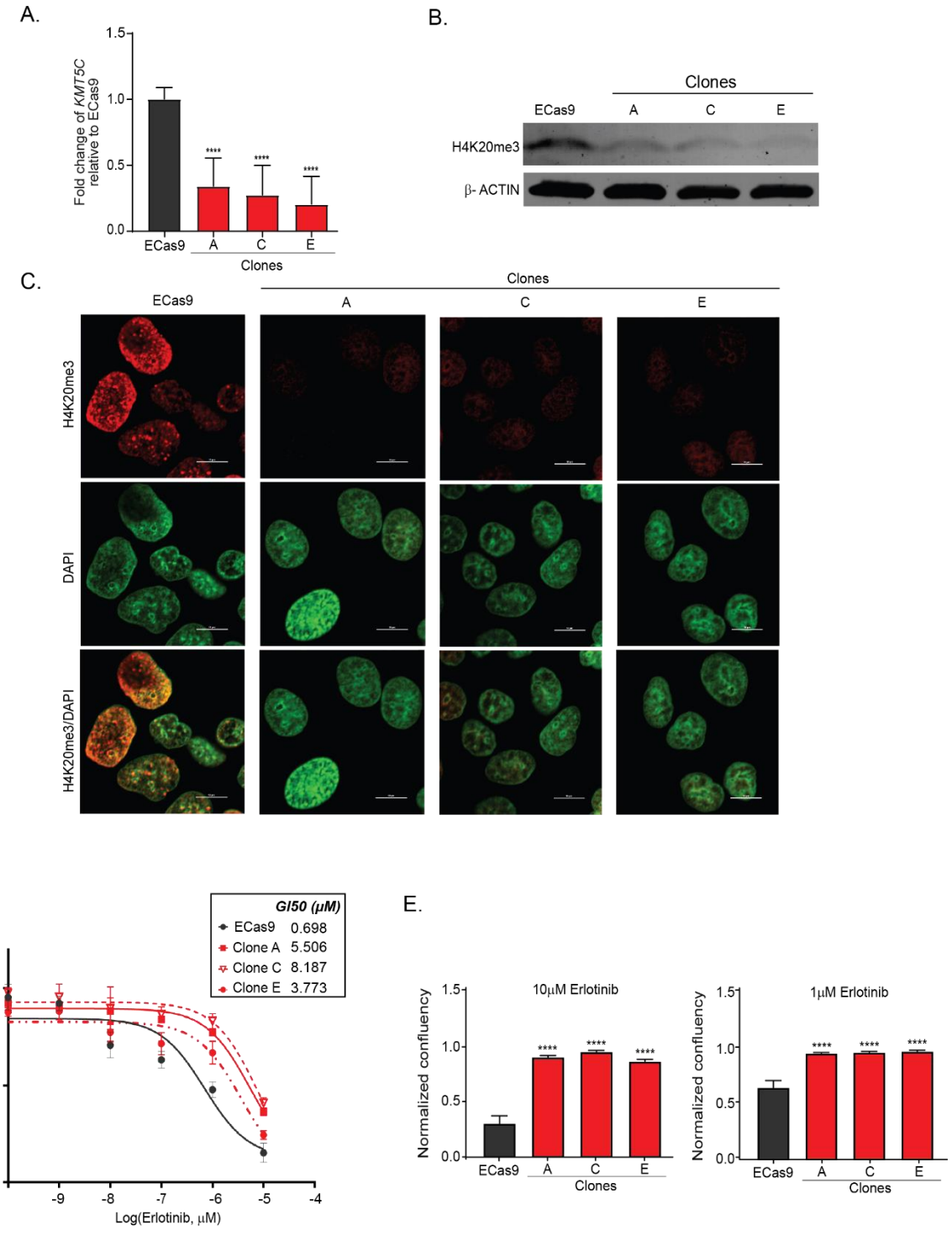
To further validate the findings from the CRISPR-Cas9 screen, we transfected ECas9 cells with a sgRNA targeting *KMT5C* to generate three SUV420H2 mutant lines, clones A, C and E. Genotyping the clones for their respective mutations validated that the sgRNA specifically targeted

KMT5C resulting in various insertions and deletions at the loci (**Supplementary Figure 2.3A**). *KMT5C* levels were reduced in the three isolated clones (**Figure 2.3A**) resulting in downregulation of H4K20me3 (**Figure 2.3B, C** and **Supplementary Figure 2.3B**). Finally, comparing erlotinib response of ECas9 cells to that of the mutant clones A, C and E confirmed that loss of SUV420H2 leads to erlotinib resistance (**Figure 2.3D**). Increased proliferation of the clones relative to ECas9 cells in the presence of erlotinib corroborated the results (**Figure 2.3E**). We also evaluated the response of SUV420H2 mutant clones to other TKIs that specifically target EGFR such as afatinib, gefitinib, and osimertinib. All the clones were resistant to all three TKIs (**Supplementary Figures 2.3C, E, G** and **Supplementary Figure 2.3D, F, H**). Conversely, loss of SUV420H2 did not confer resistance to the global DNA damaging chemotherapeutic drug cisplatin suggesting that loss of SUV420H2 does not sensitize cells to all therapeutic agents (**Supplementary Figure 2.3I**).

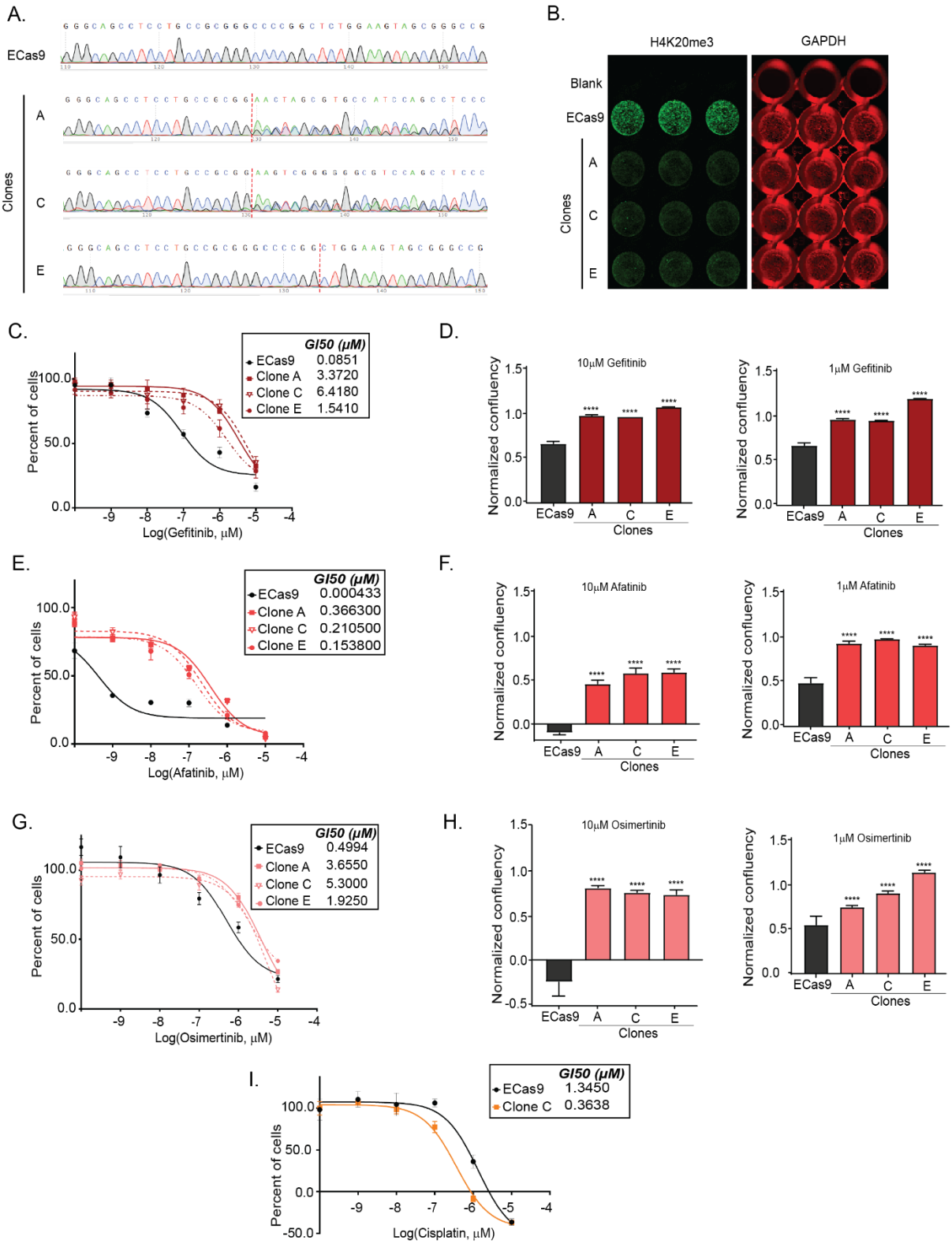
2.3.4 Ectopic expression of SUV420H2 partially sensitizes EGFR- TKI resistant cells

Since loss of SUV420H2 led to erlotinib resistance, we evaluated if the converse holds true by overexpressing SUV420H2. A doxycycline (DOX) inducible *KMT5C* plasmid was stably expressed in Calu6 cells, which have endogenously low levels of *KMT5C* (**Figure 2.1A**) and are resistant to erlotinib (**Figure 2.2A, B**). Culturing two clonally-derived lines in the presence of DOX resulted in a 4 to 8-fold increase of *KMT5C* relative to cells grown in PBS containing media (**Figure 2.4A**). H4K20me3 levels were also significantly increased following DOX induction in both clones, but not in Calu6 parental cells (**Figure 2.4B**). Culturing clones 1 and 2 in PBS or DOX containing media, followed by exposure to increasing concentrations of erlotinib resulted in partial sensitization of cells cultured in DOX, especially at higher doses of erlotinib (**Figure 2.4C**). Live-cell proliferation analysis in the presence of three different concentrations of erlotinib validated these findings (**Figure 2.4D**). With respect to gefitinib, afatinib and osimeritinib, SUV420H2 overexpressing clones were sensitized (**Supplementary Figure 2.4**), most notably at higher concentrations of each drug. Evaluating additional chemotherapeutic agents, both traditional and targeted are required to confirm that gain-of-function of SUV420H2 has a role in sensitizing resistant cells to chemotherapeutic agents.

Figure 2.3. Loss of SUV420H2 confers resistance to erlotinib. A) Endogenous expression of *KMT5C* in mutant lines, clones A, C, E evaluated by qRT-PCR. GAPDH was used as the endogenous control. One-way ANOVA was used to evaluate statistical significance of *KMT5C* transcript levels relative to ECas9 cells. B) Representative western blot image of H4K20me3 levels in ECas9 cells and clones A, C, E. β-ACTIN was used as a loading control. C) Representative immunofluorescent image of H4K20me3 in ECas9 cells and clones A, C, E. Scale bar, 10μm. D) Erlotinib dose response assayed using the SRB assay after exposing the indicated cells to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. Following normalization, GI50 concentration of erlotinib was calculated from their respective dose curves for each cell line. E) Live cell imaging or ECas9 or clones A, C, E was conducted to quantify proliferating cells in the presence of erlotinib or vehicle control (DMSO) for 72 hours. Data relative to respective normalized DMSO control treatments is represented. One-way ANOVA followed by Dunnett's Multiple Comparison test was used to evaluate statistical significance of clones A, C, E in the presence of 10 or 1μM erlotinib compared to ECas9 cells.


A.
Fold change of KMT5C relative to ECas9
1.5
1.0
0.5
0.0
****
ECas9
A
C
E
Clones
B.
Clones
ECas9
A
C
E
H4K20me3
β- ACTIN
C.
Clones
ECas9
A
C
E
H4K20me3
DAPI
H4K20me3/DAPI
D.
GI50 (μM)
ECas9 0.698
Clone A 5.506
Clone C 8.187
Clone E 3.773
Percent of cells
100.0
50.0
0.0
-9
-8
-7
-6
-5
-4
Log(Erlotinib, μM)
E.
10μM Erlotinib
1μM Erlotinib
Normalized confluency
1.5
1.0
0.5
0.0
****
ECas9
A
C
E
Clones

Supplementary Figure 2.3. SUV420H2 knock-out confers resistance to tyrosine kinase inhibitors but not to DNA damaging agents. A) Genomic DNA of ECas9 cells or clones A, C, E was isolated, the region targeted by CRISPR-Cas9 sgRNA targeting SUV420H2 was PCR amplified, purified and sequenced. Representative chromatograms of the wildtype SUV420H2 in ECas9 and the specific mutations identified in clones A, C, E. B) ICW of H4K20me3 levels in ECas9 cells and clones A, C, E. GAPDH serves as an endogenous control. C) Gefitinib, E) afatinib, G) osimertinib, or I) cisplatin dose response curves. Cells were exposed to the indicated concentration of drug or to the highest equivalent volume of vehicle control containing media for 72 hours. Following normalization, GI50concentration of each inhibitor was calculated from their respective dose curves for each cell line. Proliferation of ECas9 or clones A, C, E was evaluated using the Incucyte. Cells were exposed to varying concentrations of D) Gefitinib F) Afatinib or H) Osimertinib or the highest equivalent volume of DMSO containing media for 72 hours. Data relative to respective normalized DMSO control treatments is represented. One-way ANOVA followed by Dunnett's Multiple Comparison test was utilized to evaluate statistical significance of normalized confluency of clones A, C, E in the presence of 10 or 1µM of gefitinib, afatinib or osimertinib compared to ECas9 cells.


A.
ECas9
Clones
A
C
E
B.
H4K20me3
GAPDH
Blank
ECas9
Clones
A
C
E
C.
GI50 (μM)
ECas9 0.0851
Clone A 3.3720
Clone C 6.4180
Clone E 1.5410
Percent of cells
Log(Gefitinib, μM)
D.
10μM Gefitinib
1μM Gefitinib
Normalized confluency
ECas9 A C E
Clones
E.
GI50 (μM)
ECas9 0.000433
Clone A 0.366300
Clone C 0.210500
Clone E 0.153800
Percent of cells
Log(Afatinib, μM)
F.
10μM Afatinib
1μM Afatinib
Normalized confluency
ECas9 A C E
Clones
G.
GI50 (μM)
ECas9 0.4994
Clone A 3.6550
Clone C 5.3000
Clone E 1.9250
Percent of cells
Log(Osimertinib, μM)
H.
10μM Osimertinib
1μM Osimertinib
Normalized confluency
ECas9 A C E
Clones
I.
GI50 (μM)
ECas9 1.3450
Clone C 0.3638
Percent of cells
Log(Cisplatin, μM)

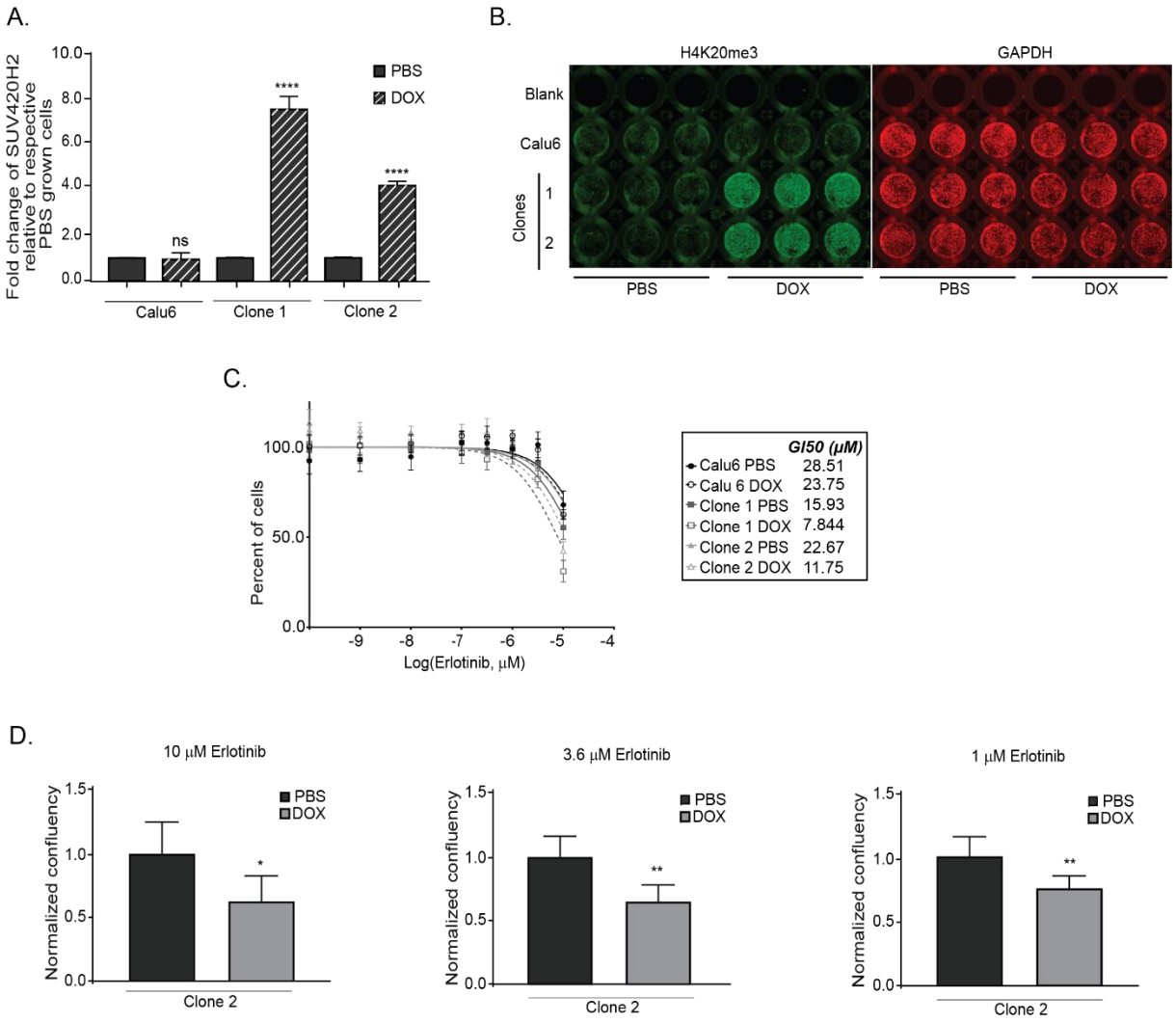
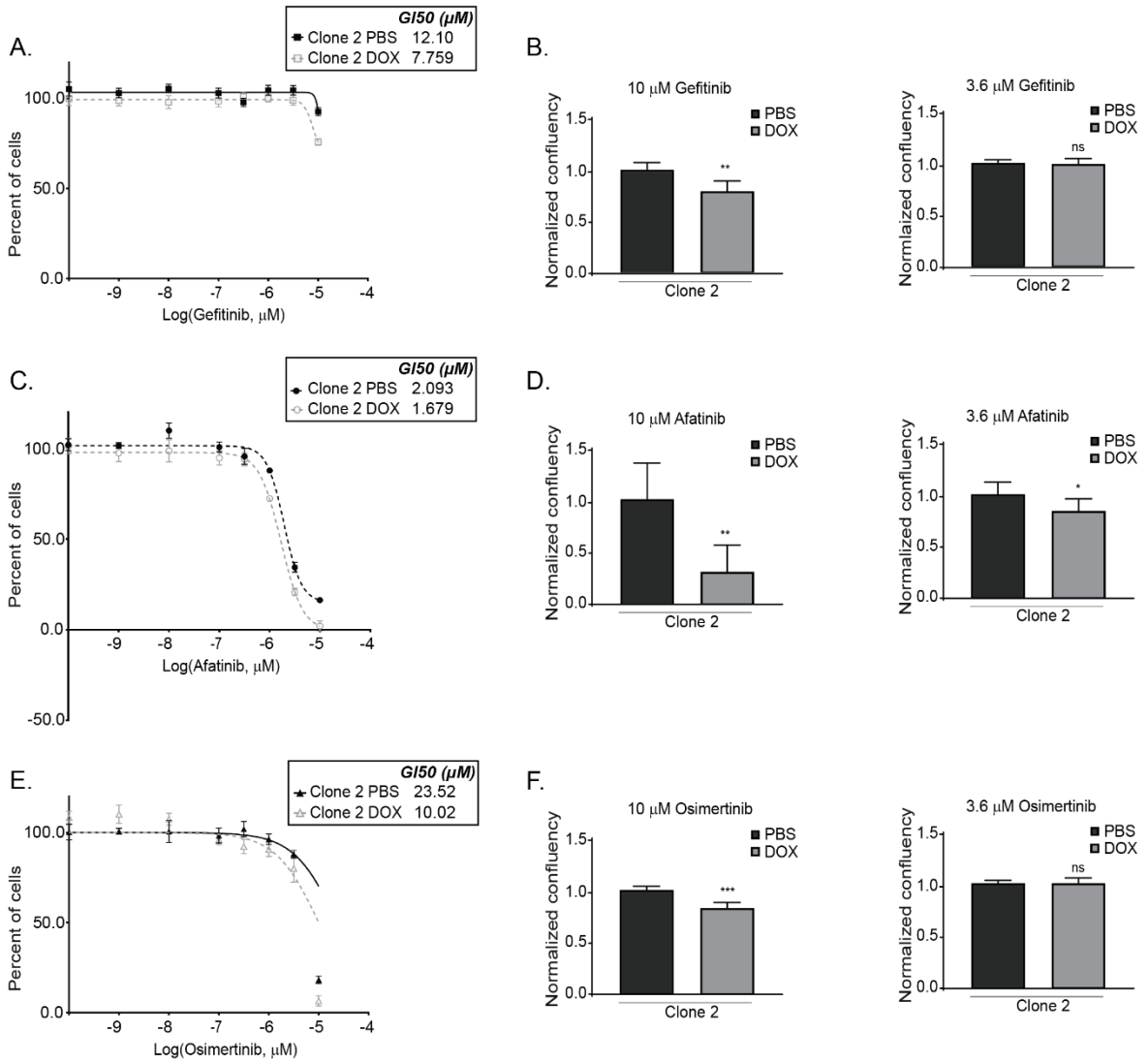


Figure 2.4. Ectopic expression of SUV420H2 partially sensitizes TKI resistant cells to erlotinib. A) SUV420H2 transcript levels evaluated by qRT-PCR in Calu6 cells and Calu6 clones 1, 2 stably expressing DOX-inducible SUV420H2. One-way ANOVA followed by Dunnett's Multiple Comparison test was utilized to evaluate statistical significance of SUV420H2 transcript levels relative to respective PBS containing media grown cells. B) H4K20me3 levels evaluated by ICW. DOX (or PBX control) treatment was for two weeks. GAPDH is used as the endogenous control. C) Erlotinib dose response measured by SRB was evaluated after a two week exposure to PBS or DOX containing media. Erlotinib treatments were for 72 hours. Following normalization, GI50 concentration of erlotinib was calculated from their respective dose curves for each cell line. D) Proliferation of clone 2 was evaluated using the Incucyte. Cells grown in PBS or DOX containing media for two weeks, were exposed to varying concentrations of erlotinib or the highest equivalent volume of DMSO containing media for 72 hours. Normalized data relative to respective normalized PBS treated samples is represented. Unpaired t-test was used to evaluate the statistical significance for each pair.



Supplementary Figure 2.4. Ectopic expression of SUV420H2 partially sensitizes resistant cells to TKIs. Dose response measured by SRB was evaluated after a two week exposure to PBS or DOX containing media for Calu6 or clones 1, 2 for A) gefitinib , C) afatininb, E) osimertinib. Erlotinib treatments were for 72 hours. Following normalization, GI50 concentration of erlotinib was calculated from their respective dose curves for each cell line. Proliferation of clone 2 was evaluated using the Incucyte. Cells grown in PBS or DOX containing media for two weeks, were exposed to varying concentrations of B) gefitinib, D) afatinib, F) osimertinib, or the highest equivalent volume of DMSO containing media for 72 hours. Unpaired t-test is utilized to evaluate statistical significance of normalized confluency of DOX containing media grown clone 2 cells in the presence of either 10, 3.6 or 1μM of gefitinib or afatinib or osimertinib compared to respective PBS containing media grown cells.

2.3.5 SUV420H2 negatively regulates the oncogenic long non-coding RNA, LINC01510 and the oncogene, MET

MET amplification is one of the more common mechanisms that cells use to circumvent blockage of EGFR signaling by erlotinib. Because SUV420H2 functions as a tumor suppressor, and is associated with repression of oncogenes (48,61), we investigated if loss of SUV420H2 alters the expression MET. We observed that MET was significantly higher in all three SUV420H2 mutant clones relative to ECas9 cells (**Figure 2.5A)** that appeared to be through a transcriptional mechanism based on elevated transcript levels (**Figure 2.5B**). Conversely, induction of SUV420H2 in clones 1, 2 resulted in reduction in MET levels (**Figure 2.5C, D**).

Previous studies determined that MET can be induced through both genomic amplification and transcriptional upregulation. Although there are multiple mechanisms that are involved in regulating transcription from the *MET* locus, a recent study identified a long non-coding RNA (lncRNA) that functions as an enhancer of *MET*. A short variant of the long non-coding RNA, LINC01510, referred to as COMET (Correlated-to-MET) (62), was identified as an enhancer of *MET* in colorectal cancer (63). High LINC01510 was also found to correlate with poor prognosis in various cancers, including NSCLC (63–65). Therefore, we analyzed the transcript levels of *LINC01510* in NSCLC, and evaluated if *LINC01510* is a negative prognostic predictor of NSCLC. GEPIA analysis in LUAD (TCGA) indicated that *LINC01510* was higher in a subset of tumors relative to normal tissues (**Supplementary Figure 2.5A**). Additionally, TANRIC analysis showed that overall survival is better in LUAD when LINC01510 transcript levels are low (**Supplementary Figure 2.5B**).

Since *LINC01510* and *MET* levels are reported to positively correlate in colorectal cancer (63), their correlation was evaluated in NSCLC. Correlation analyses using TCGA LUAD and LUSC datasets via GEPIA suggested that *LINC01510* and *MET* levels are significantly positively correlated in both LUAD (Pearson r= 0.42) and LUSC (Pearson r= 0.22) (**Figure 2.5E, F**). Finally, based on the positive correlation between *MET* and *LINC01510*, and because *LINC01510* has been reported to function as an enhancer of MET, we evaluated if SUV420H2 activity alters the expression of *LINC01510*. As hypothesized, in SUV420H2 mutant clones *LINC01510* is significantly upregulated between 8 and 10-fold (**Figure 2.5G**). Conversely, in the SUV420H2

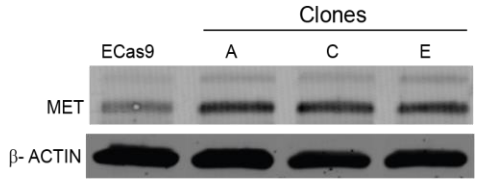
inducible clone, *LINC01510* is significantly lower when cells are cultured in the presence of DOX (**Figure 2.5H**).

2.3.6 SUV420H2 negatively regulates *LINC01510* via H4K20me3

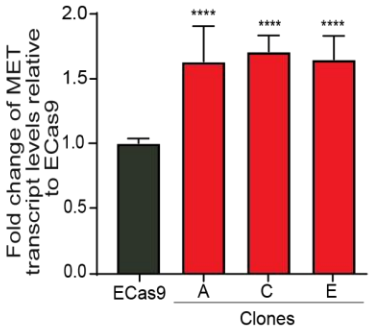
Since SUV420H2 mediates its repressive effects on oncogenes via the H4K20me3 modification (48), we hypothesized that *MET* and/or *LINC01510*, are likely negatively regulated by SUV420H2 via H4K20me3 mediated repression. To this end, we analyzed the reported H4K20me3 profile obtained via ChIP-seq of a human lung fibroblast cell line, IMR90 (GSE59316) (61). We determined that the H4K20me3 modification is present in the gene body of *LINC01510*, i.e. ~55kb upstream of start site of its neighboring gene, *MET* (**Figure 2.6A**). To identify the region of the chromosome associated with the H4K20me3 modification chromatin immunoprecipitation followed by q-RT-PCR (ChIP-qPCR) was conducted. To establish sensitivity of the assay, primers were designed and used to evaluate FOXA1, a target previously reported to be regulated by SUV420H2 via the H4K20me3 modification (66). Two primer sets were tested, one that was based on the original publication (66), and another that we designed that overlaps with the predicted H4K20me3 mark (**Supplementary Figure 2.6A and Table 2.2**). As expected, pulldown of the FOXA1 region depended on the presence of SUV420H2. A significant reduction in pulldown was observed in the SUV420H2 mutant clones (**Supplementary Figure 2.6B**) and an increase in pulldown was evident when SUV420H2 was induced (**Supplementary Figure 2.6C**).

Figure 2.5. SUV420H2 regulates the oncogene, MET and the oncogenic long non-coding RNA, LINC01510. A) Representative western blot of H4K20me3 levels or B) qRT-PCR data for MET transcript levels in ECas9 cells and SUV420H2 mutant clones. β-ACTIN was utilized as a loading control for western analysis, GAPDH serves as an endogenous control for qRT-PCR. C) Representative western blot image of H4K20me3 levels or D) qRT-PCR of *MET* transcript levels in Calu6 cells and clones stably expressing a DOX-inducible SUV420H2 vector. Cells were cultured in DOX containing media (or PBX) for two weeks. Correlation analysis between *LINC01510* and *MET* transcript levels in TCGA E) LUAD and F) LUSC datasets, evaluated using GEPIA. Expression of *LINC01510* in G) SUV420H2 mutant lines and H) in SUV420H2 inducible clones evaluated by qRT-PCR. For panels showing statistical significance, One-way ANOVA followed by Dunnett's Multiple Comparison test was used. TPM= Transcripts per million.

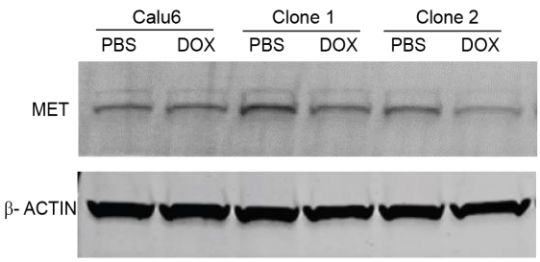
A.


Clones
ECas9
A
C
E
MET
β- ACTIN

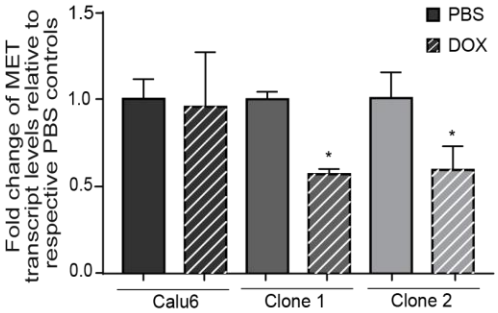
B.


Fold change of MET transcript levels relative to ECas9
2.0
1.5
1.0
0.5
0.0
****
****
****
ECas9
A
C
E
Clones

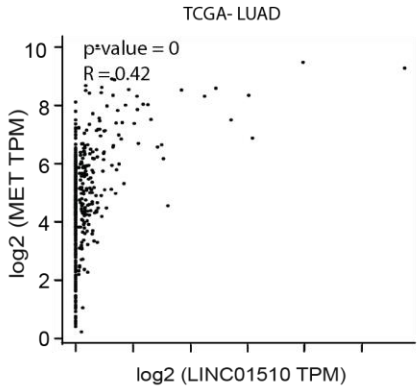
C.


Calu6
Clone 1
Clone 2
PBS
DOX
PBS
DOX
PBS
DOX
MET
β- ACTIN

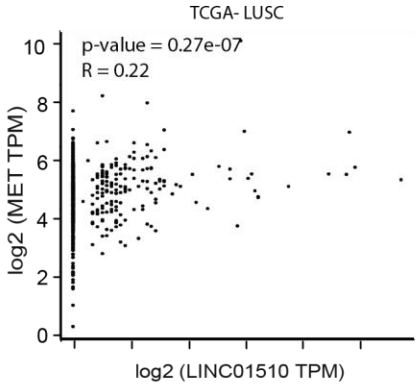
D.


Fold change of MET transcript levels relative to respective PBS controls
1.5
1.0
0.5
0.0
PBS
DOX
*
*
Calu6
Clone 1
Clone 2

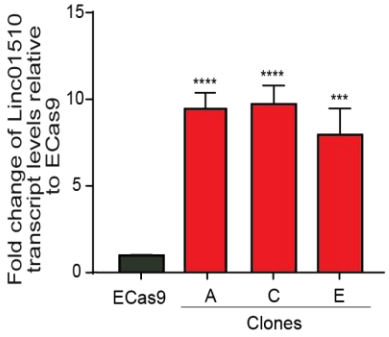
E.


TCGA- LUAD
p-value = 0
R = 0.42
log2 (MET TPM)
10
8
6
4
2
0
log2 (LINC01510 TPM)

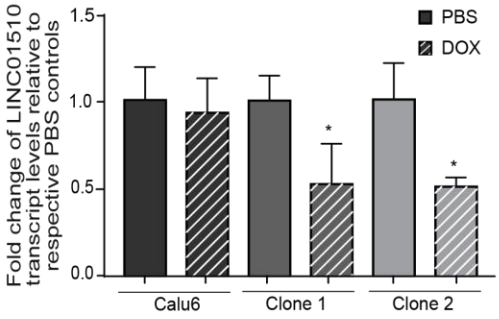
F.

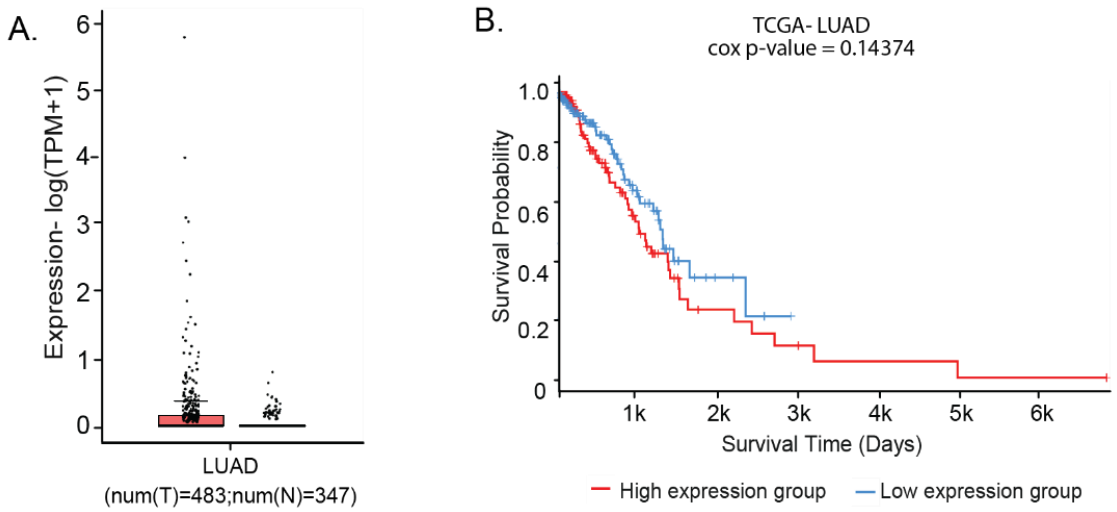

TCGA- LUSC
p-value = 0.27e-07
R = 0.22
log2 (MET TPM)
10
8
6
4
2
0
log2 (LINC01510 TPM)

G.


Fold change of Linc01510 transcript levels relative to ECas9
15
10
5
0
****
****
***
ECas9
A
C
E
Clones

H.


Fold change of LINC01510 transcript levels relative to respective PBS controls
1.5
1.0
0.5
0.0
PBS
DOX
*
*
Calu6
Clone 1
Clone 2

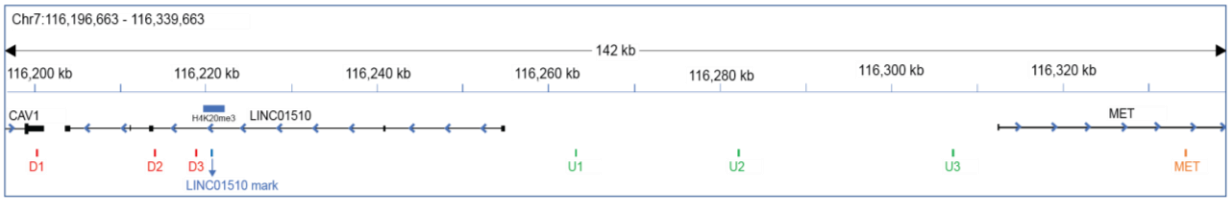


Supplementary Figure 2.5. LINC01510 correlates poorly with LUAD prognosis. A) GEPIA analysis for LINC01510 transcript levels in normal and tumor samples form LUAD data obtained from TCGA and GTEx databases. Transcript levels of *LINC01510* (ENSG00000231210.2) associated with overall survival in B) LUAD samples obtained from TCGA, and analyzed by TANRIC. TPM= Transcripts per million, T= Tumor, N=Normal.

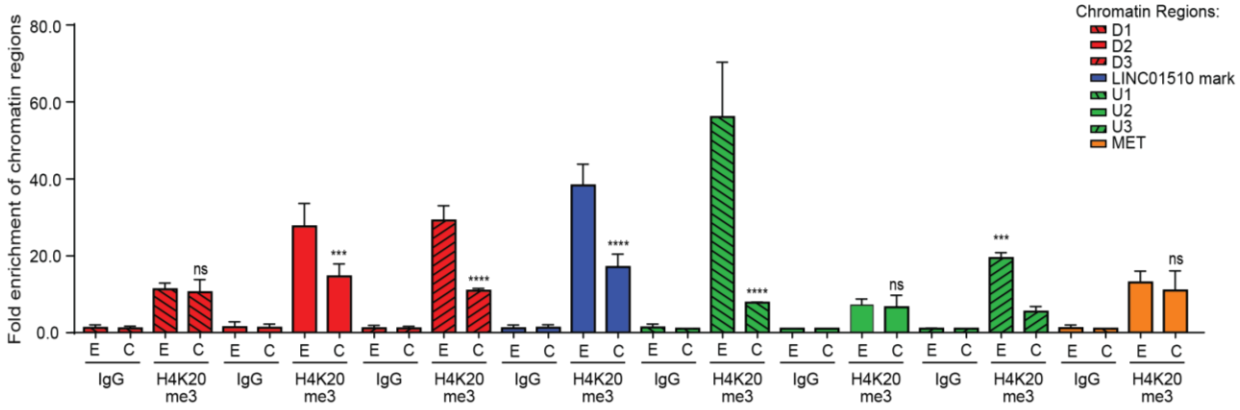
Following the results obtained from ChIP-qPCR for FOXA1, ChIP-qPCR analyses on *LINC01510* and the *MET* loci was conducted using primers that overlapped with the predicted H4K20me3 site and with primers both up and downstream (**Figure 2.6A**, **Table 2.2**). Similar to the FOXA1 locus, pulldown varied depending on the status of SUV420H2. The most abundant reduction in pull down in the SUV420H2 mutant occurred just upstream of the *LINC01510* locus with no obvious difference at the *MET* locus (**Figure 2.6B** compare upstream primer 1 (U1) to MET primers). In concordance, induction of SUV420H2 followed by ChIP-qPCR resulted in enrichment of the H4K20me3 mark in regions surrounding the lncRNA, with the LINC01510 mark having the most significant increase, with only a marginal increase at the *MET* locus (**Figure 2.6C**, compare LINC01510 primers to MET primers). Importantly the observed enrichment in ChIP-qRT-PCR for the LINC01510 mark is analogous to enrichment of the FOXA1 mark locus. These results validate that SUV420H2 regulates LINC01510 expression via the H4K20me3 mark present on its gene body.

Figure 2.6. SUV420H2 negatively regulates LINC01510 via the H4K20me3 mark present on its gene body. A) Diagram of the genomic region (Chr7: 116,196,663 - 116,339,663) representing the H4K20me3 on the *LINC01510* gene body, upstream of *MET*, as identified from GSE59316. ChIP-qPCR primers designed on and around the H4K20me3 mark. Location of ChIP-qPCR primers are indicated as LINC01510 mark, regions downstream (D1, D2, D3) and upstream (U1, U2, U3) of the H4K20me3 mark, and on MET. B) ChIP was performed on chromatin isolated from ECas9 or SUV420H2 mutant clone C using either IgG or H4K20me3 primary antibodies. E= ECas9 cells, C= SUV420H2 mutant clone C. C) ChIP was performed on chromatin isolated from DOX inducible SUV420H2 cells following growth in DOX (D, induced) or PBS (P, uninduced), using either IgG or H4K20me3 primary antibodies. qPCR using the immunoprecipitated chromatin was conducted using primers from A. Data is represented as fold enrichment of chromatin region pull-down by H4K20me3 primary antibody relative to IgG. Statistical significance is represented for fold enrichment of each chromatin region relative to its respective control. Statistical significance was conducted using one-way ANOVA followed by Dunnett's Multiple Comparison test.

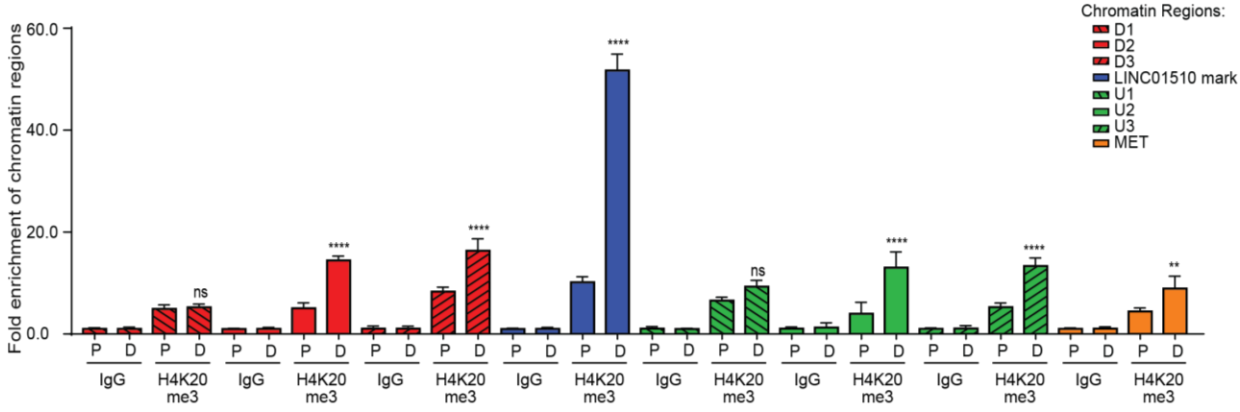
A.

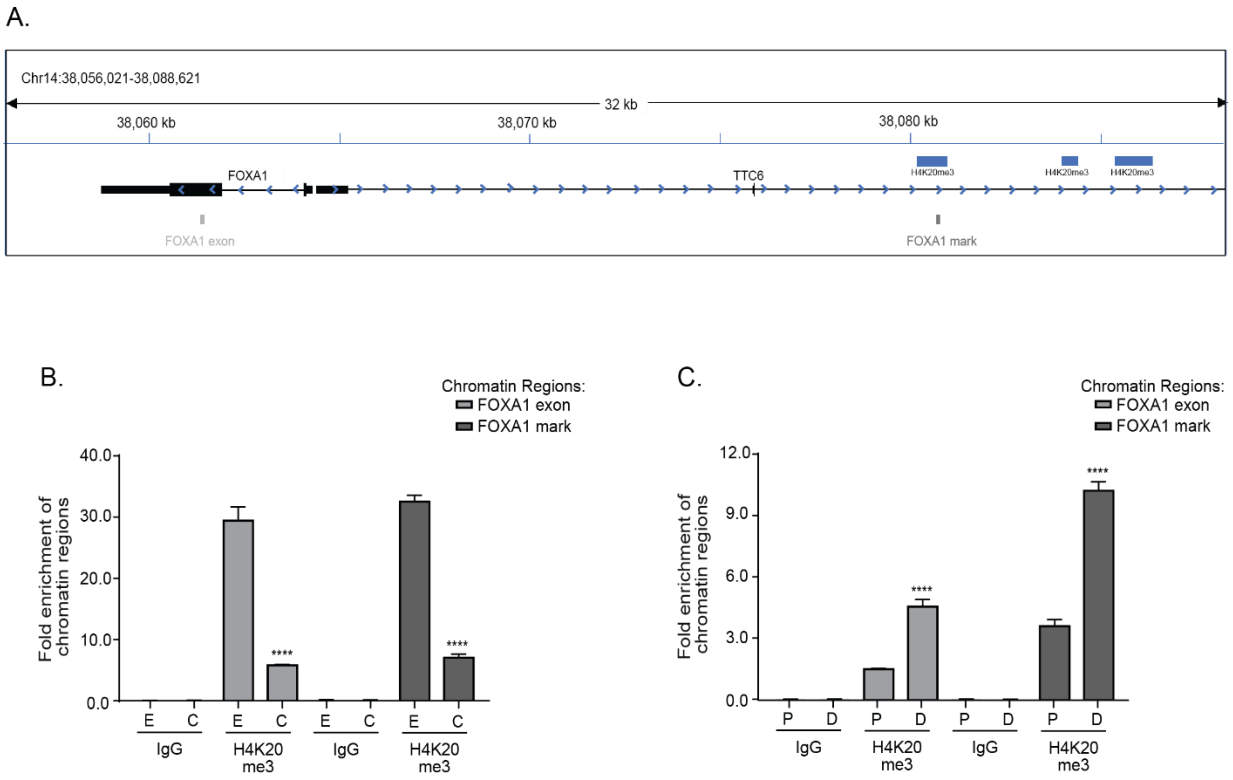

Chr7:116,196,663 - 116,339,663
142 kb
116,200 kb
116,220 kb
116,240 kb
116,260 kb
116,280 kb
116,300 kb
116,320 kb
CAV1
H4K20me3
LINC01510
MET
D1
D2
D3
LINC01510 mark
U1
U2
U3
MET

B.


Chromatin Regions:
D1
D2
D3
LINC01510 mark
U1
U2
U3
MET
Fold enrichment of chromatin regions
80.0
60.0
40.0
20.0
0.0
E
C
IgG
H4K20 me3
ns
***
****

C.


Chromatin Regions:
D1
D2
D3
LINC01510 mark
U1
U2
U3
MET
Fold enrichment of chromatin regions
60.0
40.0
20.0
0.0
P
D
IgG
H4K20 me3
ns
**
****



Supplementary Figure 2.6. FOXA1 is regulated by SUV420H2, via the H4K20me3 mark. A) ChIP-qPCR primers designed to evaluate the ChIP enrichment of the FOXA1 exonic region (FOXA1 exon), and a H4K20me3 mark upstream of the FOXA1 promoter region (FOXA1 mark), predicted through analysis of GSE59316. ChIP was performed using either IgG or H4K20me3 primary antibodies on chromatin isolated from B) ECas9 or SUV420H2 mutant clone C or C) inducible SUV420H2 cells (in the presence of DOX or PBS). qPCR using the immunoprecipitated chromatin was conducted using primers from A. Data are represented as fold enrichment of chromatin region pull-down by H4K20me3 primary antibody relative to IgG and was evaluated for significance using one-way ANOVA. E= ECas9 cells, C= SUV420H2 mutant clone C cells, P= Calu6 clones grown in PBS containing media, D= Calu6 clones grown in DOX containing media.

2.3.7 Knockdown of *LINC01510* or *MET* partially re-sensitize SUV420H2 mutant cells to erlotinib, via downregulation of MET

From figure 5, it can be inferred that SUV420H2 negatively regulates both *LINC01510* and *MET* levels, and MET protein levels, whereas figure 6 validated that SUV420H2, via H4K20me3 negatively regulates *LINC01510* at the transcriptional level. Therefore, we evaluated if SUV420H2 indirectly negatively regulates MET via *LINC01510* downregulation, which is a reported enhancer of *MET* transcription (62,63). *LINC01510* or *MET* were knocked down in the SUV420H2 mutant clone, which expresses high levels of *LINC01510* and *MET* (**Figures 2.5A, B, G**). By western blot

and qRT-PCR analyses, it was confirmed that siRNAs targeting either *MET* or *LINC01510* downregulate MET at both the transcript and protein level (**Figure 2.7A-D**). To determine if loss of SUV420H2 partially mediates erlotinib resistance via upregulation of *LINC01510* and *MET*, *LINC01510* or *MET* were downregulated and erlotinib dose response and proliferation analyses were conducted. Both results validate that erlotinib resistant SUV420H2 mutant cells can be partially re-sensitized to erlotinib post knockdown of either *LINC01510* or *MET* (**Figure 2.7E, F**).

2.3.8 Overexpression of *LINC01510* or MET leads to development of erlotinib resistance in sensitive cells, via upregulation of MET

Data presented in figure 7 suggests that LINC01510 is a *bona fide* enhancer of MET, therefore, to further evaluate if *LINC01510* is a trans-acting enhancer, *LINC01510* was overexpressed in parental ECas9 cells and MET levels were quantified. Following transfection of the *LINC01510* plasmid, and *MET* plasmids individually, a modest increase in MET was observed (**Figure 2.8A-D**). Additionally, as hypothesized, *LINC01510* and *MET* overexpression also led to acquired resistance in ECas9 cells, evaluated by both dose curve and proliferation analyses (**Figure 2.8E, 8F**).

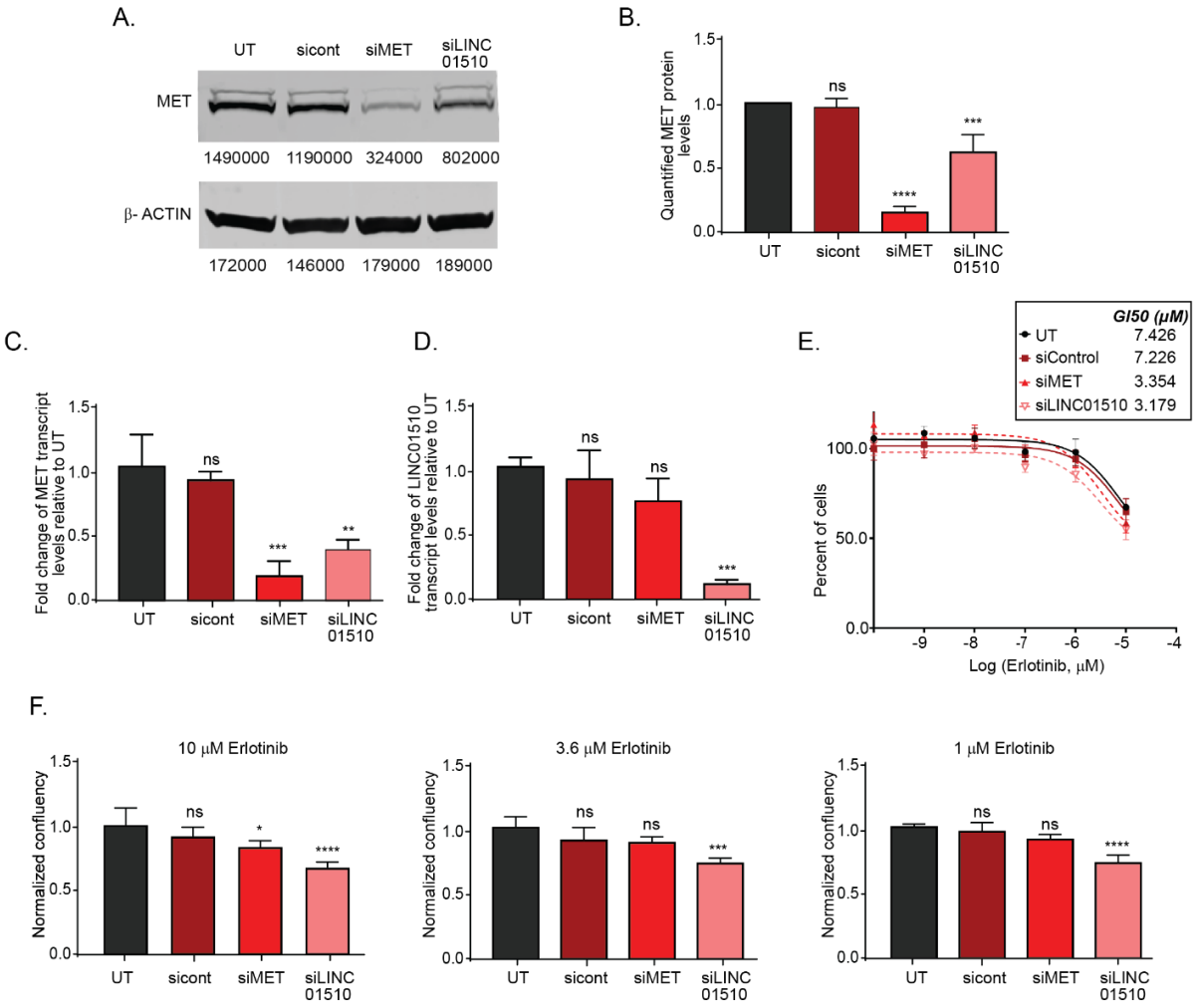


Figure 2.7. Knockdown of LINC01510 or MET partially re-sensitizes SUV420H2 mutant cells to erlotinib, via downregulation of MET. A) Representative western blot of MET in SUV420H2 mutant cells that were either untransfected (UT) or reverse transfected with siRNA control (sicont), siRNA to MET (siMET), or siRNA to LINC01510 (siLINC01510) for 96 hours. β-ACTIN serves as a loading control. B) Quantification of protein levels from three biological replicates as done in A. Expression of C) *MET* and D) *LINC01510* in SUV420H2 mutant cells that were either UT or reverse transfected with sicont, siMET or siLINC01510 for 96 hours. GAPDH is utilized as the endogenous control. E) erlotinib dose response of SUV420H2 mutant cells following transfection with the indicted siRNAs. Twenty-four hours post transfection, cells were exposed to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. Post-normalization, GI50 concentration of erlotinib was calculated from each respective dose curves for each transfection. E) Proliferation of SUV420H2 mutant cells following transfection with the indicted siRNAs. Twenty-four hours post transfection, cells were exposed to varying concentrations of erlotinib or the highest equivalent volume of DMSO containing media for 72 hours. Normalized data is represented relative to that of UT. One-way ANOVA followed by Dunnett's Multiple Comparison test was used to evaluate statistical significance.

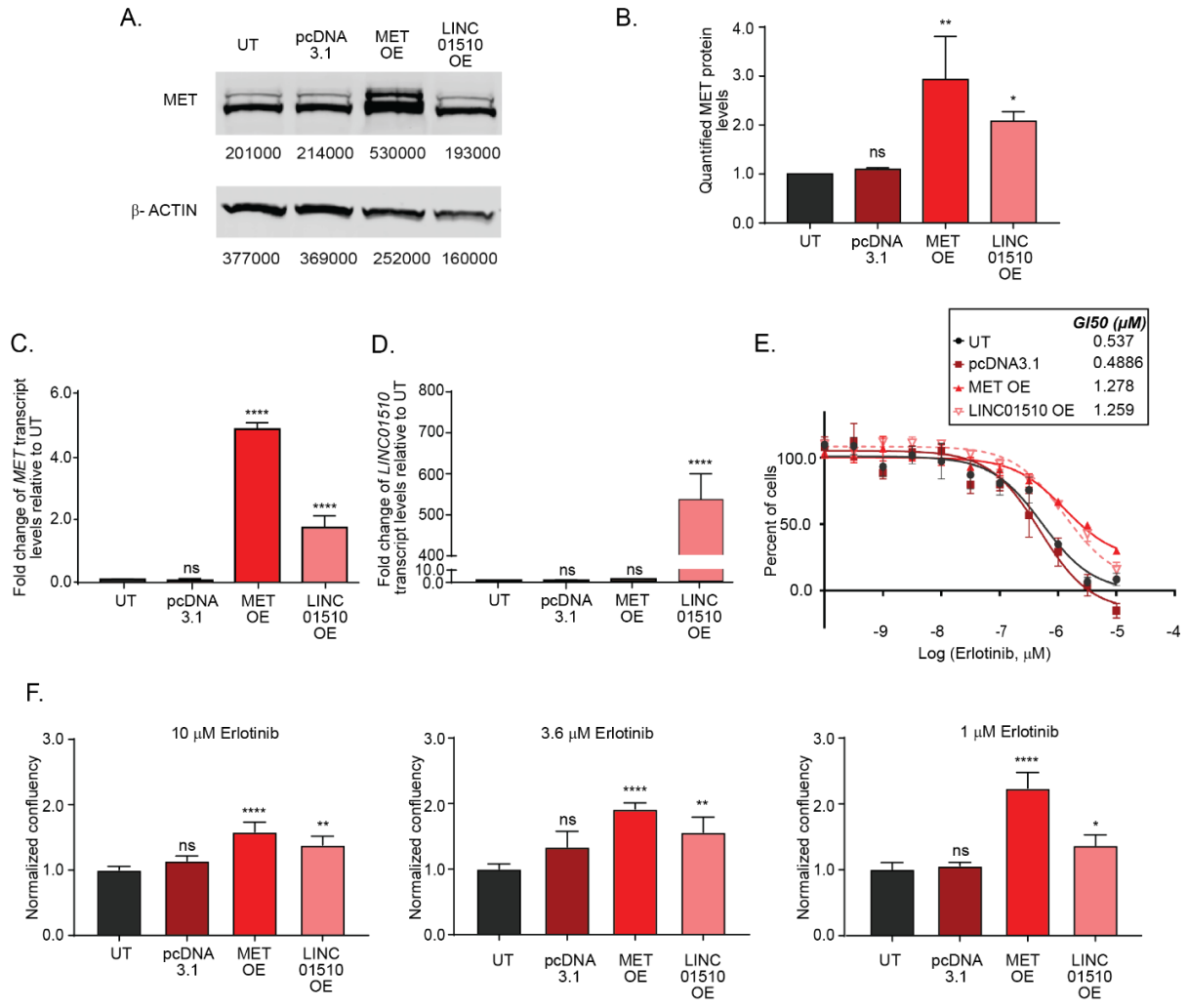


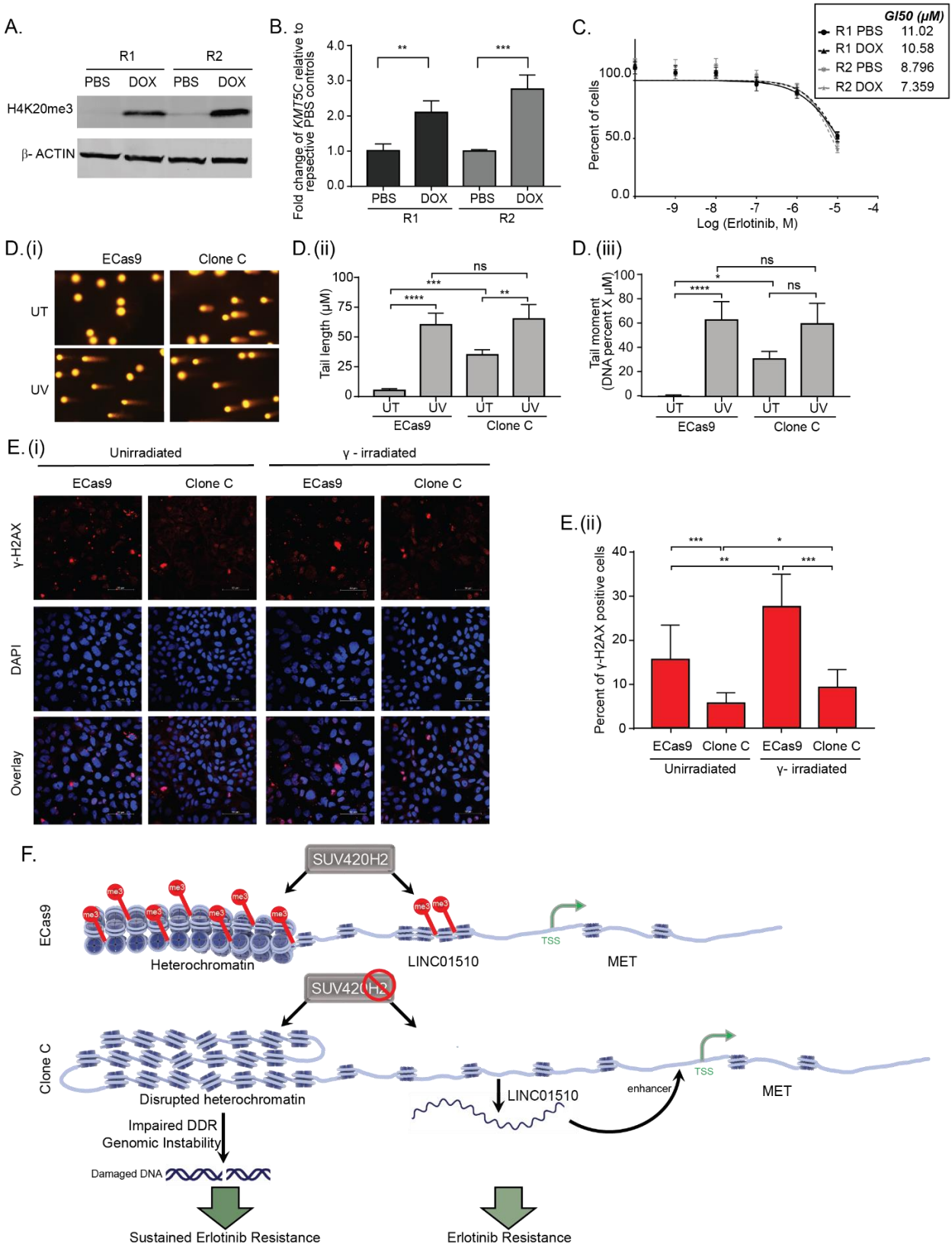
Figure 2.8. Overexpression of LINC01510 or MET promotes erlotinib resistance, via enhanced expression of MET A) Representative western blot of MET in ECsa9 cells that were either untransfected (UT) or transfected with pcDNA3.1 control plasmid or plasmids to overexpress to MET (MET OE) or LINC01510 (LINC01510 OE) for 96 hours. β-ACTIN was used as a loading control. B) Quantification of MET levels from three biological replicates as in A. Expression of C) *MET* and D) *LINC01510* in ECas9 cells that were either UT or transfected with the indicated vectors. Data are normalized to *GAPDH*. E) Erlotinib dose response via SRB assay was evaluated in ECas9 cells that were either UT or that were transfected with the indicated vectors. Twenty-four hours post transfection, cells were exposed to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. Post-normalization, GI50 concentration of erlotinib was calculated from each respective dose curves for each transfection. F) Proliferation of ECas9 cells transfected as in E. Twenty-four hours post transfection, cells were exposed to varying concentrations of erlotinib or the highest equivalent volume of DMSO containing media for 72 hours. Normalized data is represented relative to that of UT. One-way ANOVA followed by Dunnett's Multiple Comparison test was used to evaluate statistical significance.

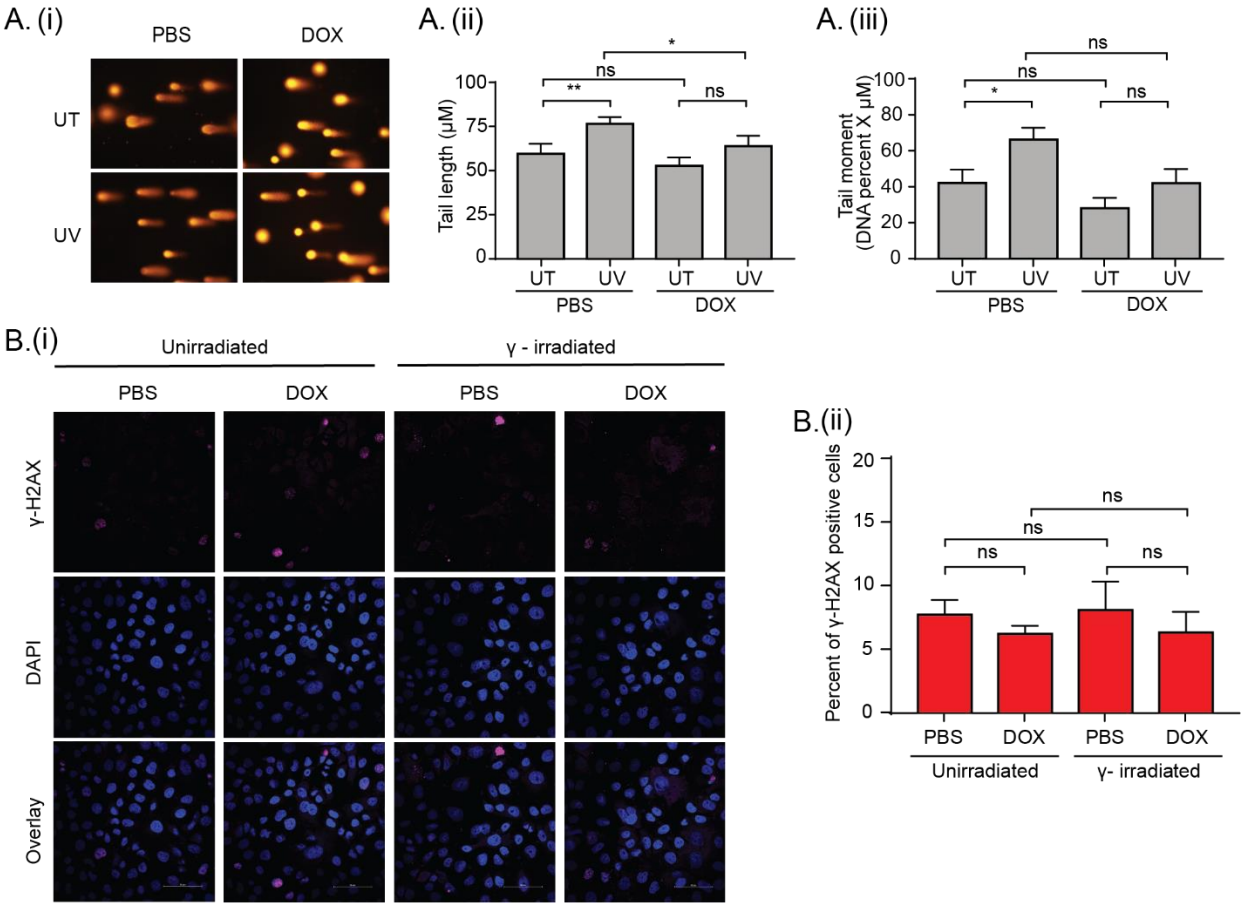
2.3.9 SUV420H2 mutant cells are genomically unstable, preventing rescue of erlotinib resistant phenotype upon SUV420H2 re-expression

As loss of SUV420H2 promoted resistance to TKIs, we hypothesized that rescuing the expression of SUV420H2 in the mutant cells may re-sensitize the cells. SUV420H2 mutant cells were stably transfected with an SUV420H2-inducible plasmid to generate DOX-inducible rescue lines. Following culturing the rescue lines (R1, R2) in DOX containing media for one month, H4K20me3 and *KMT5C* were both upregulated (**Figure 2.9A, B**), however, neither of the clones were re-sensitized to erlotinib (**Figure 2.9C**). Therefore, we evaluated if genomic instability, reported to occur in SUV420H double knock-out mice (25), may be responsible for the inability of cells to revert the TKI resistance phenotype following re-expression of SUV420H2. Genomic DNA fragmentation was analyzed in both the SUV420H2 mutant clone and in ECas9 cells via comet assays. Data suggest that loss of SUV420H2 causes genomic instability (**Figure 2.9D**), and rescuing SUV420H2 expression in the mutant cells failed to re-stabilize genomic integrity (**Supplementary Figure 2.7A**). Further analysis of γ-H2AX, a histone variant involved in the DNA damage repair (DDR) pathway, determined that SUV420H2 mutant cells have lower baseline γ-H2AX relative to ECas9 cells. When cells were exposed to DNA damaging gamma (γ) radiation, 11.9 % of ECas9 cells had upregulation of γ-H2AX, while only 3.6% of SUV420H2 mutant cells induced γ-H2AX (**Figure 2.9E**). Even in the presence of a strong DNA damaging agent, the SUV420H2 mutant cells had ~40% less γ-H2AX levels than unirradiated ECas9 cells. Whereas for PBS or DOX exposed rescue clones, R1 and R2, no significant difference was observed in γ-H2AX levels post γ-irradiation relative to unirradiated cells (**Supplementary Figure 2.7B**).

Overall, the model in Figure 7G depicts that loss of SUV420H2 in NSCLC cells results in de-repression of the oncogenic lncRNA, LINC01510 that mediates erlotinib resistance by enhancing transcription of the oncogene MET. Additionally, SUV420H2 mutant cells display a dampened DDR mechanism, therefore accumulating damaged DNA and genomic instability, sustaining erlotinib resistance in the cells (**Figure 2.7F**).

Figure 2.9. Mutant SUV420H2 cells display sustained erlotinib resistance via accumulation of genomic instability. A) Representative western blot of H4K20me3 levels in rescue clones (R1, R2) generated by stably expressing an SUV420H2-inducible plasmid in SUV420H2 mutant cells. Cells were grown for one month in media containing PBS or DOX. B) Expression of *KMT5C* in rescue clones R1 and R2 cells, grown in PBS or DOX containing media. C) Erlotinib dose response via SRB assay was evaluated following culturing cells in either PBS or DOX and varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. Post-normalization, the GI50 concentration of erlotinib was calculated from each respective dose curve for each transfection. D) (i) Comet assays were performed on 10,000 ECas9 or mutant cells (Clone C). Cells were either unexposed to UV (UT) or exposed to 900 J/m^2 of UV. (ii) Tail length (µM) or (iii) Tail moment (DNA percent X tail length) was evaluated using OpenComet plugin via ImageJ. E) (i) Representative IF image of γ-H2AX in WT and mutant clone c. Cells were either unirradiated or irradiated with 4Gy of γ-radiation. DAPI was used as an endogenous control. IF of γ-H2AX overlapped with DAPI is represented in the third panel. Scale bar, 50µm. (ii) Cells that are positive for γ-H2AX relative to the total number of cells (DAPI positive cells) was quantified from n = 2 biological replicates, represented as percentage. F) Model depicting a mechanism by which loss of SUV420H2 mediates erlotinib resistance in NSCLC cells (right). The model also illustrates a mechanism by which SUV420H2 mutant cells incur genomic instability and fail to recover from erlotinib resistance.


A.
R1
R2
PBS
DOX
PBS
DOX
H4K20me3
β- ACTIN
B.
Fold change of KMT5C relative to repsective PBS controls
4.0
3.0
2.0
1.0
0.0
**
***
PBS
DOX
PBS
DOX
R1
R2
C.
GI50 (μM)
R1 PBS 11.02
R1 DOX 10.58
R2 PBS 8.796
R2 DOX 7.359
Percent of cells
100.0
50.0
0.0
-9
-8
-7
-6
-5
-4
Log (Erlotinib, M)
D.(i)
ECas9
Clone C
UT
UV
D.(ii)
Tail length (μM)
100
75
50
25
0
ns
***
****
**
UT
UV
UT
UV
ECas9
Clone C
D. (iii)
Tail moment (DNA percent X μM)
100
80
60
40
20
0
ns
*
****
ns
UT
UV
UT
UV
ECas9
Clone C
E.(i)
Unirradiated
γ - irradiated
ECas9
Clone C
ECas9
Clone C
γ-H2AX
DAPI
Overlay
E.(ii)
Percent of γ-H2AX positive cells
40
30
20
10
0
***
*
**
***
ECas9
Clone C
ECas9
Clone C
Unirradiated
γ- irradiated
F.
SUV420H2
ECas9
me3
Heterochromatin
LINC01510
TSS
MET
SUV420H2
Clone C
Disrupted heterochromatin
LINC01510
enhancer
TSS
MET
Impaired DDR
Genomic Instability
Damaged DNA
Sustained Erlotinib Resistance
Erlotinib Resistance



Supplementary Figure 2.7. SUV420H2 re-expression in mutant cells does not rescue genomic instability of cells. A) (i)A Comet assay was performed on 10,000 rescue cells (R2) that were grown for one month in media containing PBS or DOX. Cells were either unexposed (UT) or exposed to 900 J/m^2 of UV. (ii) Tail length (μM) or (iii) Tail moment (DNA percent X tail length) was evaluated. B) (i) Representative IF image of γ-H2AX in PBS or DOX exposed R2 cells. Cells were either unirradiated or irradiated with 4Gy of γ-radiation. DAPI was utilized as an endogenous control. IF of γ-H2AX overlapped with DAPI is represented in the third panel. Scale bar, 50μm. (ii) Cells that are positive for γ-H2AX relative to the total number of cells (DAPI positive cells) was quantified, represented as percentage.

2.4 Discussion:

Changes to the epigenome are common occurrences that influence all aspects of cancer (Review 61), including chemoresistance. However, only a limited subset of epigenetic factors have been determined to have a role in resistance to therapeutic drugs in cancer (Reviews 62,63). The aim of this study was to identify novel mechanisms by which acquired erlotinib resistance manifests in NSCLC, and loss of SUV420H2 was identified as one such novel epigenetic factor. SUV420H2 is a histone methyltransferase responsible for maintaining heterochromatic states or repressing specific regions of the genome such as oncogenes, via the repressive mark H4K20me3. Both SUV420H2 and H4K20me3 are significant for maintaining cells in their differentiated states, loss of which is consequentially reported to cause enhanced survival due to elongation of telomeres (70,71), stemness of cancer cells (72), and spontaneous carcinogenesis (28,29,58,73).

Alongside, inability to establish H4K20me3 via dysregulated SUV420H2 is reported to result in less efficient repair of damaged DNA by the DNA damage repair (DDR) mechanism leading to severe genomic instability (24,25,74–77). Analogously, loss of SET8, a methyltransferase that catalyzes H4K20 monomethylation serving as a substrate for SUV420H2 enzymatic activity, leads to decompaction of heterochromatic states. These compromised chromatin regions become vulnerable to damage (76,78,79). I*n vivo* studies have also validated that SUV420H1 and H2 double knock-out mice are perinatally lethal. Such mice have lost the majority of their H4K20me di- (catalyzed by SUV420H1) and trimethylation (catalyzed by SUV420H2) marks and have acquired a genome-wide H4K20 monomethylation state (25). This genome-wide transition allows for accumulation of rapid DNA damage due to impaired double-strand break (DSB) repair resulting in severe chromosomal aberrations. Therefore, SUV420H2 functions as an epigenetic factor responsible for protecting and stabilizing regions of the chromatin, which when misregulated can lead to the generation of multiple cancerous phenotypes.

The results of this study describe that loss of SUV420H2 confers erlotinib resistance to NSCLC cells via pleotropic effects. One novel mechanism defined is that loss of SUV420H2 abrogates the H4K20me3 modification at an oncogenic long non-coding RNA, *LINC01510*. Enhanced transcription of *LINC01510* which is reported to function as a transcriptional enhancer of the oncogene *MET* consequently leads to *MET* upregulation, a predominant mechanism of acquired

resistance to erlotinib. Therefore, this study establishes a mechanism of erlotinib resistance mediated by loss of SUV420H2, which in part is due to indirect overexpression of *MET*.

Loss of SUV420H2 also led to genomic instability, which has not been commonly reported as a mechanism of EGFR-TKI resistance (80,81). However, it has been long appreciated that genomic instability generates tumor heterogeneity and in the presence of a drug, gives rise to resistant cells (Reviews 62,63,77). In the current study, loss of functional SUV420H2 may have led to spontaneous genetic aberrations leading to rapid establishment of resistant population of cells in the presence of erlotinib. Moreover, induction of SUV420H2 in Calu6 cells which have endogenously modest levels of *KMT5C*, and high levels of H4K20me3 likely due to the presence of an additional H4K20 trimethylating enzyme (83,84), cells were partially sensitized to TKIs. Analogously, rescue clones generated from mutated SUV420H2 cells (clone c), did not re-sensitize to erlotinib upon one week of SUV420H2 induction, but did show a partial reversion of resistance phenotype after a month of SUV420H2 induction (*data collection in process*). These findings suggest that maintenance of a basal level of H4K20me3 modification of the chromatin, allows for a partial reversion of resistance in NSCLC cells. It is possible that in EKVX (parental cells of clone c), SUV420H2 is the major H4K20 trimethylating enzyme, therefore in the absence of SUV420H2 the chromatin may have suffered massive loss of H4K20me3, which was acting as a heterochromatic shield to prevent DNA damage.

Previous reports also determined that loss of SUV420H1/2 impair the DDR mechanism, inadvertently leading to accumulation of damaged DNA and increased tumorigenicity (74,85,86). These findings, and our data suggests a possible direct involvement of SUV420H2 and H4K20me3 in the recruitment of mediators of the DDR pathway in NSCLC cells, loss of which results in acquired erlotinib resistance.

Although the findings of this study, for the first time show that SUV420H2 is a mediator of drug resistance, loss of SUV39H1/2, an upstream regulator of SUV420H2 has previously been reported to be associated with resistance (87,88). In corroboration with the genomic instability observed in SUV420H2 mutant cells in our study, it has been reported that SUV39H null mice displayed chromosomal instabilities and increased tumorigenicity (76,89). Additionally, the first ever identified demethylase for H4K20me3, mdig was determined to be overexpressed in breast and

lung cancer cells antagonizing the effects of H4K20me3 mark, which led to induction of expression of oncogenes (90). Analogous to mdig overexpression in cancer cells, leading to reduction of H4K20me3, we found that loss of SUV420H2 also leads to depletion of H4K20me3 mark that in turn enhances oncogenes such as LINC01510 and MET.

In conclusion, it can be inferred that SUV420H2, which is a genome-wide facilitator of the epigenetic mark H4K20me3 is required to maintain genomic stability and repression of genes, such as oncogenes and oncogenic lncRNAs. Loss of SUV420H2 causes pleotropic effects such as spontaneous genomic instability and upregulation of oncogenes, mediating erlotinib resistance in NSCLC cells.

2.5 Limitations of the study:

Since this study is aimed at understanding how loss of SUV420H2 mediates erlotinib resistance in NSCLC, it would have been ideal to use antibodies to evaluate protein levels of SUV420H2. However, due to lack of a specific SUV420H2 antibody, we have utilized antibodies that detect H4K20me3, and performed qRT-PCR to evaluate *KMT5C* levels. Also, having *in vivo* and/or patient data, pre- and post- erlotinib treatment would solidify the findings of this study.

2.6 References:

1. Lung cancer atlas. The Lung Cancer https://canceratlas.cancer.org/the-burden/lung-cancer/ [Internet]. [cited 2019 Dec 22]. Available from: https://canceratlas.cancer.org/the-burden/lung-cancer/
2. Luo B, Wing H, Subramanian A, Sharifnia T, Okamoto M, Yang X, et al. Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci. 2008;105:20380–20385.
3. Hirsch FR, Varella-Garcia M, Bunn PA, Di Maria M V., Veve R, Bremnes RM, et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: Correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol. 2003;21:3798–807.

4. Onitsuka T, Uramoto H, Ono K, Takenoyama M, Hanagiri T, Oyama T, et al. Comprehensive molecular analyses of lung adenocarcinoma with regard to the epidermal growth factor Receptor, K-ras, MET, and hepatocyte growth factor status. J Thorac Oncol. Lippincott Williams and Wilkins; 2010;5:591–6.
5. Howe LR, Leevers SJ, Gómez N, Nakielny S, Cohen P, Marshall CJ. Activation of the MAP kinase pathway by the protein kinase raf. Cell [Internet]. 1992 [cited 2020 Feb 2];71:335–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1330321
6. Riely GJ, Marks J, Pao W. KRAS mutations in non-small cell lung cancer. Proc Am Thorac Soc. 2009. page 201–5.
7. Lee JW, Soung YH, Kim SY, Nam SW, Park WS, Wang YP, et al. ERBB2 kinase domain mutation in the lung squamous cell carcinoma. Cancer Lett. 2006;237:89–94.
8. Hirai F, Edagawa M, Shimamatsu S, Toyozawa R, Toyokawa G, Nosaki K, et al. Evaluation of erlotinib for the treatment of patients with non-small cell lung cancer with epidermal growth factor receptor wild type. Oncol Lett. Spandidos Publications; 2017;14:306–12.
9. Inno A, Di Noia V, Martini M, D'Argento E, Di Salvatore M, Arena V, et al. Erlotinib for Patients with EGFR Wild-Type Metastatic NSCLC: a Retrospective Biomarkers Analysis. Pathol Oncol Res. Springer Netherlands; 2019;25:513–20.
10. Osarogiagbon RU, Cappuzzo F, Ciuleanu T, Leon L, Klughammer B. Erlotinib therapy after initial platinum doublet therapy in patients with EGFR wild type non-small cell lung cancer: Results of a combined patient-level analysis of the NCIC CTG BR.21 and SATURN trials. Transl Lung Cancer Res. AME Publishing Company; 2015;4:465–74.
11. Shepherd FA, Pereira JR, Ciuleanu T, Eng HT, Hirsh V, Thongprasert S, et al. Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med [Internet]. 2005 [cited 2020 Jan 9];353:123–32. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa050753
12. Camidge DR, Pao W, Sequist L V. Acquired resistance to TKIs in solid tumours: Learning from lung cancer. Nat. Rev. Clin. Oncol. Nature Publishing Group; 2014. page 473–81.
13. Liao BC, Griesing S, Yang JCH. Second-line treatment of EGFR T790M-negative non-small cell lung cancer patients. Ther. Adv. Med. Oncol. SAGE Publications Inc.; 2019. page 1–16.

14. Gambacorti-Passerini CB, Gunby RH, Piazza R, Galietta A, Rostagno R, Scapozza L. Molecular mechanisms of resistance to imatinib in Philadelphia-chromosome-positive leukaemias. Lancet Oncol [Internet]. 2003 [cited 2020 Feb 2];4:75–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12573349
15. Yun CH, Mengwasser KE, Toms A V., Woo MS, Greulich H, Wong KK, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc Natl Acad Sci U S A. 2008;105:2070–5.
16. Niederst MJ, Sequist L V., Poirier JT, Mermel CH, Lockerman EL, Garcia AR, et al. RB loss in resistant EGFR mutant lung adenocarcinomas that transform to small-cell lung cancer. Nat Commun. Nature Publishing Group; 2015;6:6377.
17. Sequist L V., Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med. 2011;3:75ra26.
18. Shien K, Toyooka S, Yamamoto H, Soh J, Jida M, Thu KL, et al. Acquired resistance to EGFR inhibitors is associated with a manifestation of stem cell-like properties in cancer cells. Cancer Res. 2013;73:3051–61.
19. Jakobsen KR, Demuth C, Madsen AT, Hussmann D, Vad-Nielsen J, Nielsen AL, et al. MET amplification and epithelial-to-mesenchymal transition exist as parallel resistance mechanisms in erlotinib-resistant, EGFR-mutated, NSCLC HCC827 cells. Oncogenesis. Nature Publishing Group; 2017;6.
20. Sos ML, Koker M, Weir BA, Heynck S, Rabinovsky R, Zander T, et al. PTEN loss contributes to erlotinib resistance in EGFR-mutant lung cancer by activation of akt and EGFR. Cancer Res. 2009;69:3256–61.
21. Huang S, Benavente S, Armstrong EA, Li C, Wheeler DL, Harari PM. P53 modulates acquired resistance to EGFR inhibitors and radiation. Cancer Res. 2011;71:7071–9.
22. Forloni M, Gupta R, Nagarajan A, Politi K, Dogra SK. Oncogenic EGFR Represses the TET1 DNA Demethylase to Induce Silencing of Tumor Suppressors in Cancer Cells. Cell Rep [Internet]. 2016 [cited 2020 Jan 6];16:457–71. Available from: http://dx.doi.org/10.1016/j.celrep.2016.05.087

23. de Bruin EC, Cowell C, Warne PH, Jiang M, Saunders RE, Melnick MA, et al. Reduced NF1 expression confers resistance to EGFR inhibition in lung cancer. Cancer Discov. 2014;4:606–19.
24. Hahn M, Dambacher S, Dulev S, Kuznetsova AY, Eck S, Wörz S, et al. Suv4-20h2 mediates chromatin compaction and is important for cohesion recruitment to heterochromatin. Genes Dev. 2013;27:859–72.
25. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callén E, et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev. Cold Spring Harbor Laboratory Press; 2008;22:2048–61.
26. Weirich S, Kudithipudi S, Jeltsch A. Specificity of the SUV4-20H1 and SUV4-20H2 protein lysine methyltransferases and methylation of novel substrates. J Mol Biol [Internet]. Elsevier Ltd; 2016;428:2344–58. Available from: http://dx.doi.org/10.1016/j.jmb.2016.04.015
27. Bosch-Presegué L, Raurell-Vila H, Thackray JK, González J, Casal C, Kane-Goldsmith N, et al. Mammalian HP1 Isoforms Have Specific Roles in Heterochromatin Structure and Organization. Cell Rep. 2017;21:2048–57.
28. Pogribny IP, Tryndyak VP, Muskhelishvili L, Rusyn I, Ross SA. Methyl Deficiency, Alterations in Global Histone Modifications, and Carcinogenesis 1,2 [Internet]. J. Nutr. Int. Res. Conf. Food. 2007. Available from: https://academic.oup.com/jn/article-abstract/137/1/216S/4664374
29. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37:391–400.
30. Golden RJ, Chen B, Li T, Braun J, Manjunath H, Chen X, et al. An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. Nature [Internet]. Nature Publishing Group; 2017;542:197–202. Available from: http://dx.doi.org/10.1038/nature21025
31. Li W, Köster J, Xu H, Chen CH, Xiao T, Liu JS, et al. Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. Genome Biol [Internet]. Genome Biology; 2015;16:1–13. Available from: http://dx.doi.org/10.1186/s13059-015-0843-6

32. Liu XS, Zhang F, Irizarry RA, Xiao T, Cong L, Love MI, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 2014;15:1–12.
33. LentiGuide-Puro and LentiCRISPRv2. http://www.genome-engineering.org/gecko/. [cited 2020 Jan 12]; Available from: http://www.genome-engineering.org/gecko/
34. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res [Internet]. 2017 [cited 2020 Jan 12];45:W98–102. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkx247
35. Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, et al. TANRIC: An interactive open platform to explore the function of IncRNAs in cancer. Cancer Res. American Association for Cancer Research Inc.; 2015;75:3728–37.
36. NCI-60 DTP. NCI-60 Screening Methodology | NCI-60 Human Tumor Cell Lines Screen | Discovery & Development Services | Developmental Therapeutics Program (DTP) https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm [Internet]. [cited 2019 Dec 30]. Available from: https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm
37. Orellana E, Kasinski A. Sulforhodamine B (SRB) Assay in Cell Culture to Investigate Cell Proliferation. BIO-PROTOCOL. Bio-Protocol, LLC; 2016;6:e1984.
38. Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. Nat Rev Genet [Internet]. 2015;16:299–311. Available from: http://dx.doi.org/10.1038/nrg3899
39. Chen D, Si W, Shen J, Du C, Lou W, Bao C, et al. MiR-27b-3p inhibits proliferation and potentially reverses multi-chemoresistance by targeting CBLB/GRB2 in breast cancer cells. Cell Death Dis [Internet]. Springer US; 2018;9. Available from: http://dx.doi.org/10.1038/s41419-017-0211-4
40. Chen S, Wang Q, Zhou XM, Zhu JP, Li T, Huang M. MicroRNA-27b reverses docetaxel resistance of non-small cell lung carcinoma cells via targeting epithelial growth factor receptor. Mol Med Rep. 2016;14:949–54.

41. Paul I, Chacko AD, Stasik I, Busacca S, Crawford N, McCoy F, et al. Acquired differential regulation of caspase-8 in cisplatin-resistant non-small-cell lung cancer. Cell Death Dis. 2012;3.
42. Ferreira CG, Span SW, Peters GJ, Kruyt FAE, Giaccone G. Chemotherapy Triggers Apoptosis in a Caspase-8-dependent and Mitochondria-controlled Manner in the Non-Small Cell Lung Cancer Cell Line NCI-H460. Cancer Res [Internet]. American Association for Cancer Research; 2000;60:7133–41. Available from: https://cancerres.aacrjournals.org/content/60/24/7133
43. Krammer PH, Mattern J, Herr I, Ucur E, Debatin K-M, Okouoyo S, et al. Rescue of death receptor and mitochondrial apoptosis signaling in resistant human NSCLCin vivo. Int J Cancer. 2004;108:580–7.
44. Zhang G, Scarborough H, Kim J, Rozhok AI, Chen YA, Zhang X, et al. Coupling an EML4-ALK-centric interactome with RNA interference identifies sensitizers to ALK inhibitors. Sci Signal [Internet]. American Association for the Advancement of Science; 2016 [cited 2020 Jan 9];9:rs12–rs12. Available from: https://stke.sciencemag.org/lookup/doi/10.1126/scisignal.aaf5011
45. Orzáez M, Guevara T, Sancho M, Pérez-Payá E. Intrinsic caspase-8 activation mediates sensitization of erlotinib-resistant tumor cells to erlotinib/cell-cycle inhibitors combination treatment. Cell Death Dis. 2012;3:1–9.
46. CtRP v2. Cancer Therapeutics Response Portal https://portals.broadinstitute.org/ctrp.v2.1/ [Internet]. [cited 2020 Jan 9]. Available from: https://portals.broadinstitute.org/ctrp.v2.1/
47. Zhou H, Liu Y, Xiao L, Hu Z, Xia K. Overexpression of MicroRNA-27b inhibits proliferation, migration, and invasion via suppression of MET expression. Oncol Res. 2017;25:147–54.
48. Shinchi Y, Hieda M, Nishioka Y, Matsumoto A, Yokoyama Y, Kimura H, et al. SUV420H2 suppresses breast cancer cell invasion through down regulation of the SH2 domain-containing focal adhesion protein tensin-3. Exp Cell Res [Internet]. Elsevier; 2015;334:90–9. Available from: http://dx.doi.org/10.1016/j.yexcr.2015.03.010
49. Aprelikova O, Palla J, Hibler B, Yu X, Greer YE, Yi M, et al. Silencing of miR-148a in cancer-associated fibroblasts results in WNT10B-mediated stimulation of tumor cell motility. Oncogene. 2013;32:3246–53.

50. Ardeshir-Larijani F, Bhateja P, Lipka MB, Sharma N, Fu P, Dowlati A. KMT2D Mutation Is Associated With Poor Prognosis in Non–Small-Cell Lung Cancer. Clin Lung Cancer [Internet]. Elsevier Inc.; 2018;19:e489–501. Available from: https://doi.org/10.1016/j.cllc.2018.03.005
51. Kantidakis T, Saponaro M, Mitter R, Horswell S, Kranz A, Boeing S, et al. Mutation of cancer driver MLL2 results in transcription stress and genome instability. Genes Dev. 2016;30:408–20.
52. Chen Y, Saif Zaman M, Deng G, Majid S, Saini S, Liu J, et al. MicroRNAs 221/222 and genistein-mediated regulation of ARHI tumor suppressor gene in prostate cancer. Cancer Prev Res. 2011;4:76–86.
53. Ryu SW, Yoon J, Yim N, Choi K, Choi C. Downregulation of OPA3 Is Responsible for Transforming Growth Factor-β-Induced Mitochondrial Elongation and F-Actin Rearrangement in Retinal Pigment Epithelial ARPE-19 Cells. PLoS One. 2013;8:1–9.
54. Taylor JE, Lawson KA, Wongtawan T, Wilmut I, Pennings S. Histone H4K20me3 and HP1 are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. J Cell Sci. 2011;124:1878–90.
55. Bierhoff H, Grummt I, Schotta G, Dammert MA, Dambacher S, Brocks D. Quiescence-Induced LncRNAs Trigger H4K20 Trimethylation and Transcriptional Silencing. Mol Cell [Internet]. Elsevier Inc.; 2014;54:675–82. Available from: http://dx.doi.org/10.1016/j.molcel.2014.03.032
56. Nishioka Y, Yokoyama Y, Tsujimoto M, Kimura H, Yoshidome K, Ogihara E, et al. Loss of histone H4K20 trimethylation predicts poor prognosis in breast cancer and is associated with invasive activity. Breast Cancer Res. 2014;16.
57. Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, et al. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. Clin Cancer Res. 2008;14:7237–45.
58. Tryndyak VP, Kovalchuk O, Pogribny IP. Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone methyltransferase and methyl-binding proteins. 2006 [cited 2020 Jan 10]; Available from: https://doi.org/10.4161/cbt.5.1.2288

59. GTEx. GTEx Portal https://gtexportal.org/home/ [Internet]. [cited 2020 Feb 2]. Available from: https://gtexportal.org/home/
60. TCGA. The Cancer Genome Atlas Program https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga [Internet]. [cited 2020 Feb 2]. Available from: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga
61. Nelson DM, Jaber-Hijazi F, Cole JJ, Robertson NA, Pawlikowski JS, Norris KT, et al. Mapping H4K20me3 onto the chromatin landscape of senescent cells indicates a function in control of cell senescence and tumor suppression through preservation of genetic and epigenetic stability. Genome Biol [Internet]. 2016;17:158. Available from: https://doi.org/10.1186/s13059-016-1017-x
62. Esposito R, Esposito D, Pallante P, Fusco A, Ciccodicola A, Costa V. Oncogenic properties of the antisense lncRNA COMET in BRAF- and RET-driven papillary thyroid carcinomas. Cancer Res. American Association for Cancer Research Inc.; 2019;79:2124–35.
63. Li J, Lin C, Zhang T, Zuo Y, Liu J, Li X, et al. Long noncoding RNA LINC01510 promotes the growth of colorectal cancer cells by modulating MET expression. Cancer Cell Int [Internet]. BioMed Central; 2018;18:1–12. Available from: https://doi.org/10.1186/s12935-018-0503-5
64. Li J, Wei L. Increased expression of LINC01510 predicts poor prognosis and promotes malignant progression in human non-small cell lung cancer. Biomed Pharmacother [Internet]. Elsevier; 2019;109:519–29. Available from: https://doi.org/10.1016/j.biopha.2018.10.136
65. Li Q, Wang X jun, Jin J hong. SOX2-induced upregulation of lncRNA LINC01510 promotes papillary thyroid carcinoma progression by modulating miR-335/SHH and activating Hedgehog pathway. Biochem Biophys Res Commun. Elsevier B.V.; 2019;520:277–83.
66. Viotti M, Wilson C, McCleland M, Koeppen H, Haley B, Jhunjhunwala S, et al. SUV420H2 is an epigenetic regulator of epithelial/ mesenchymal states in pancreatic cancer. J Cell Biol. Rockefeller University Press; 2018;217:763–77.

67. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. Science (80-) [Internet]. American Association for the Advancement of Science; 2017 [cited 2020 Jan 15];357. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.aal2380
68. Wilting RH, Dannenberg JH. Epigenetic mechanisms in tumorigenesis, tumor cell heterogeneity and drug resistance. Drug Resist Updat [Internet]. Elsevier Ltd; 2012;15:21–38. Available from: http://dx.doi.org/10.1016/j.drup.2012.01.008
69. Dannenberg Jan-Hermen JH, Berns A. Drugging Drug Resistance. Cell. 2010;141:18–20.
70. Benetti R, García-Cao M, Blasco MA. Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres. Nat Genet. 2007;39:243–50.
71. Marión RM, Schotta G, Ortega S, Blasco MA. Suv4-20h abrogation enhances telomere elongation during reprogramming and confers a higher tumorigenic potential to iPs cells. PLoS One. 2011;6:e25680.
72. Liu X, Li F, Huang Q, Zhang Z, Zhou L, Deng Y, et al. Self-inflicted DNA double-strand breaks sustain tumorigenicity and stemness of cancer cells. Cell Res. Nature Publishing Group; 2017;27:764–83.
73. Bagnyukova T V, Tryndyak VP, Montgomery B, Churchwell MI, Karpf AR, James SR, et al. Genetic and epigenetic changes in rat preneoplastic liver tissue induced by 2-acetylaminofluorene. Carcinogenesis [Internet]. 2008 [cited 2020 Jan 10];29:638–46. Available from: https://academic.oup.com/carcin/article-abstract/29/3/638/2476765
74. Bromberg KD, Mitchell TRH, Upadhyay AK, Jakob CG, Jhala MA, Comess KM, et al. The SUV4-20 inhibitor A-196 verifies a role for epigenetics in genomic integrity. Nat Chem Biol. 2017;13:317–24.
75. Jørgensen S, Schotta G, Sørensen CS. Histone H4 Lysine 20 methylation: Key player in epigenetic regulation of genomic integrity. Nucleic Acids Res. 2013;41:2797–806.
76. Kovaríková AS, Legartová S, Krejcí J, Bártová E. H3K9me3 and H4K20me3 represent the epigenetic landscape for 53BP1 binding to DNA lesions. Aging (Albany NY). Impact Journals LLC; 2018;10:2585–605.
77. Sanders SL, Portoso M, Mata J, Bähler J, Allshire RC, Kouzarides T. Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell. 2004;119:603–14.

78. Jørgensen S, Elvers I, Trelle MB, Menzel T, Eskildsen M, Jensen ON, et al. The histone methyltransferase SET8 is required for S-phase progression. J Cell Biol. 2007;179:1337–45.
79. Shoaib M, Walter D, Gillespie PJ, Izard F, Fahrenkrog B, Lleres D, et al. Histone H4K20 methylation mediated chromatin compaction threshold ensures genome integrity by limiting DNA replication licensing. [cited 2020 Jan 10]; Available from: www.nature.com/naturecommunications
80. Serizawa M, Takahashi T, Yamamoto N, Koh Y. Genomic aberrations associated with erlotinib resistance in non-small cell lung cancer cells. Anticancer Res. 2013;33:5223–34.
81. Nahar R, Zhai W, Zhang T, Takano A, Khng AJ, Lee YY, et al. Elucidating the genomic architecture of Asian EGFR-mutant lung adenocarcinoma through multi-region exome sequencing. Nat Commun. Nature Publishing Group; 2018;9:216.
82. Gillies RJ, Verduzco D, Gatenby RA. Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. Nat Rev Cancer [Internet]. 2012 [cited 2020 Jan 16];12:487–93. Available from: www.nature.com/reviews/cancer
83. Dai L, Ye S, Li H-W, Chen D-F, Wang H-L, Jia S-N, et al. SETD4 Regulates Cell Quiescence and Catalyzes the Trimethylation of H4K20 during Diapause Formation in Artemia. Mol Cell Biol. American Society for Microbiology; 2017;37.
84. Ye S, Ding Y-F, Jia W-H, Liu X-L, Feng J-Y, Zhu Q, et al. Tumor Biology and Immunology SET Domain-Containing Protein 4 Epigenetically Controls Breast Cancer Stem Cell Quiescence. 2019 [cited 2020 Jan 10]; Available from: http://cancerres.aacrjournals.org/
85. Bassing CH, Suh H, Ferguson DO, Chua KF, Manis J, Eckersdorff M, et al. Histone H2AX: A dosage-dependent suppressor of oncogenic translocations and tumors. Cell. Cell Press; 2003;114:359–70.
86. Celeste A, Difilippantonio S, Difilippantonio MJ, Fernandez-Capetillo O, Pilch DR, Sedelnikova OA, et al. H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. Cell. Cell Press; 2003;114:371–83.
87. Braig M, Lee S, Loddenkemper C, Rudolph C, Peters AHFM, Schlegelberger B, et al. Oncogene-induced senescence as an initial barrier in lymphoma development. Nature. 2005;436:660–5.

88. Völkel P, Souza PP, Vandamme J, Angrand P-O, Trinel D, Héliot L, et al. The histone methyltransferase SUV420H2 and Heterochromatin Proteins HP1 interact but show different dynamic behaviours. BMC Cell Biol. 2009;10:41.
89. Peters AHFM, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C, et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell. Cell Press; 2001;107:323–37.
90. Zhang Q, Thakur C, Fu Y, Bi Z, Wadgaonkar P, Xu L, et al. Mdig promotes oncogenic gene expression through antagonizing repressive histone methylation markers. Theranostics [Internet]. 2020 [cited 2020 Jan 10];10:602–14. Available from: http://www.thno.org//creativecommons.org/licenses/by/4.0/

CHAPTER 3. IDENTIFICATION OF NOVEL MICRORNA MEDIATORS OF ERLOTINIB RESISTANCE IN NON-SMALL CELL LUNG CANCER

Parts of the following chapter has been adapted from a first author publication.
Pal, Arpita S, and Andrea L Kasinski. "Animal Models to Study MicroRNA Function." Advances in cancer research vol. 135 (2017): 53-118. doi:10.1016/bs.acr.2017.06.006
(License (License # 4744250655345) for reuse in thesis/dissertation, for both print and electronic formats are provided by Elsevier and Copyright Clearance Center)

Chapter Overview

This chapter describes the details of microRNA biogenesis and the normal cellular functions of microRNAs. Following which, the role of dysregulated microRNAs in the development and progression of cancer, especially NSCLC and the role of microRNAs as mediators of erlotinib resistance in NSCLC are touched upon. Lastly, the below study describes a comprehensive microRNA screen conducted with the hypothesis that there are additional unknown microRNAs that mediate erlotinib resistance that remain to be identified and validated.

3.1 Introduction

The central dogma of molecular biology, laid down by Francis Crick in 1958, stated that the fundamental role of RNA molecules is to transmit the genetic code into proteins(1,2). However, with the characterization of the first transfer RNA (tRNA) in 1965(3), additional RNAs emerged that violated the central dogma. For example, ribosomal RNAs (rRNAs)(4), small nuclear RNAs (snRNAs)(5) and small nucleolar RNAs (snoRNAs)(6) are not translated into protein products like a messenger RNAs (mRNAs), but indirectly influence the process of protein synthesis. Such RNA molecules were collectively termed "non-coding RNAs"(Review(7)). While the novel roles performed by some non-coding RNAs were being assimilated as imperative molecular mechanisms, a study conducted to identify genes in heterochronic signaling incidentally led to the discovery of an additional non-coding RNA, a 22 nucleotide RNA molecule, *lin-4*(8). The discovery of *lin-4* further defied the central dogma of molecular biology via an unprecedented

mechanism and led to the establishment of a new class of small non-coding RNAs called "microRNAs"(9-11).

Lin-4 was the first microRNA (miRNA) identified, which was determined to be indispensable for the normal development of *Caenorhabditis elegans*(8). Functionally, *lin-4* interacts with the 3′-untranslated region (3′-UTR) of the mRNA transcript *lin-14*, resulting in a marked repression of the lin-14 protein(12). Unfortunately, due to limited knowledge in RNA biology at the time, *lin-4* and its peculiar role were overlooked to be a worm-specific phenomenon. Seven years later, a second *C. elegans* miRNA, *let-7* was discovered which encouraged further miRNA investigations(13). It became apparent that *let-7* was not only critical for the development of *C. elegans*, but is also evolutionarily conserved in other organisms, including humans(14). Currently ~2,500 human encoded miRNAs have been identified, which are listed in a miRNA database, miRBase (http://www.mirbase.org/, Release 21)(15). In addition to their identification, biochemical and molecular studies have determined that the canonical function of miRNAs is to post-transcriptionally regulate a repertoire of protein-coding mRNA transcripts, whereas a few miRNAs perform unanticipated or "non-canonical" functions (Review(16,17)).

Following the identification of these first two miRNAs, *lin-4* and *let-7*, our understanding of miRNAs in normal physiology and diseased states, such as cancer (Review(18)) has advanced remarkably. Advancements in the field have been possible due to state-of-the-art technologies such as high throughput screening and deep sequencing, but majorly due to the development of appropriate *in vivo* model systems (Review(19)). Although the focus of this review is on various model systems that have been instrumental in elucidating the roles of miRNAs in cancers, and the technologies that have been extensively applied to generate these animal model systems, only the introduction section is incorporated in this chapter to introduce miRNAs in cancer. Briefly, the role of miRNAs in the development of various cancers, including lung cancer, and their roles in the development of resistance will also be touched upon.

3.1.1 MicroRNA biogenesis, mechanism of action and function

Biogenesis

Expression of miRNA genes:

The transcription of miRNA genes is regulated by multiple mechanisms eventually dictating the level of expression of a particular miRNA in normal or diseased states (Reviews(20-22)).

(a) Regulation mediated by availability of transcription factors:

A transcription factor can enhance or repress the expression of a miRNA gene depending on the availability of the particular factor (Reviews(23,24)). The prominent tumor-suppressor p53 which functions as a transcription factor for several genes, also enhances the transcription of miRNA genes. Examples of miRNAs that are directly induced by p53 include *mir-34a* and *b/c*(25-28) and two miR-200 subfamilies, *mir-200c/141* and *mir-200a/200b/429*(29,30). Akin to protein-coding genes, some transcription factors can repress a cohort of miRNAs while inducing the expression of other. For example, myelocytomatosis factor (MYC) a well-studied oncogenic transcription factor, negatively regulates the tumor suppressive miRNA, *let-7a-1*(31,32), while it positively regulates the oncogenic miR-17~92 cluster(33,34).

(b) Regulation mediated by genomic location of miRNA genes:

(i) Location in the epigenome:

Transcriptional activation or inactivation of specific miRNA genes is largely influenced by epigenetics. Such epigenetic regulation includes the proximity of the miRNA gene promoter to a CpG island, various histone modifications to the chromatin, and availability of factors that maintain and regulate expression from the epigenome. The expression of *mir-127* and *mi-34a*, miRNAs located near CpG islands are dependent on the methylation status of the promoter, implying epigenetic control on the expression of these miRNAs (35-39). MiRNAs under the control of DNA methylation also undergo massive upregulation when the DNA methylatransferases 1 and 3b (DNMT1, DNMT3) are downregulated(35), lending further support to the role of DNA methylation in regulating miRNA expression (Reviews(23,24)).

(ii) Location relative to host genes:

The origin of a miRNA gene from a specific chromosomal location impacts the extent of expression of the miRNA. In the context of other genes, miRNAs genes are either intragenic where they are embedded within a host gene, or intergenic if they are located between two genes on a chromosome. Expression of an intragenic miRNA is dependent on the expression of the host gene (Reviews(23,24)). MiR-126 is one such miRNA whose expression is concomitantly controlled by epigenetic regulation of its host gene *EGFL7*(40). Intragenic miRNAs are also regulated by canonical mechanisms that influence host gene expression such as transcription factor occupancy at the promoter of the host gene(28,41) (Review(42)). MiRNAs that are not directly regulated by a host gene are still subject to nearby epigenetic influence. For example, *let-7a-3* and miR-129-1 expression are dependent on a nearby region of the genome that is prone to altered methylation states during the onset of cancer. The proximity to this differentially methylated region severely impacts their expression(24,43,44).

(iii) Regulation by miRNA copy number:

A single mature miRNA can be expressed and processed from multiple loci in the genome. For example, three individual genes encoding human *mir-7* produce an identical mature miRNA product(45). Conversely, miR-21 is generated from a single genomic locus(46,47). A similar situation is observed for the miR-17~92 cluster found on chromosome 13, which has two paralogs: the miR-106b~25 cluster on chromosome 7 and the miR-106a~363 cluster on the X chromosome(48-51). The advantages of miRNAs originating from various loci relative to one originating from a single locus is discussed in a later section.

(iv) Cancer-Associated Genomic Regions (CAGRs):

Specific regions in the human genome that are prone to amplification or loss upon the onset of cancers are referred to as Cancer-Associated Genomic Regions (CAGRs). CAGRs contain amplified or deleted miRNA and/or protein-coding genes. Many of these genetic aberrations are required for tumorigenesis. MiRNAs that are lost are frequently located in either fragile sites of the genome or regions susceptible to loss of heterozygosity (LOH). For example, the *mir-15a~16-1* cluster located in a fragile region of the genome at 13q14.3 is frequently deleted in Chronic Lymphocytic Leukemia (CLL) patients(28,52,53). Whereas other miRNAs are commonly amplified in multiple cancers due to their location in fragile regions. For example, the 17q23-25

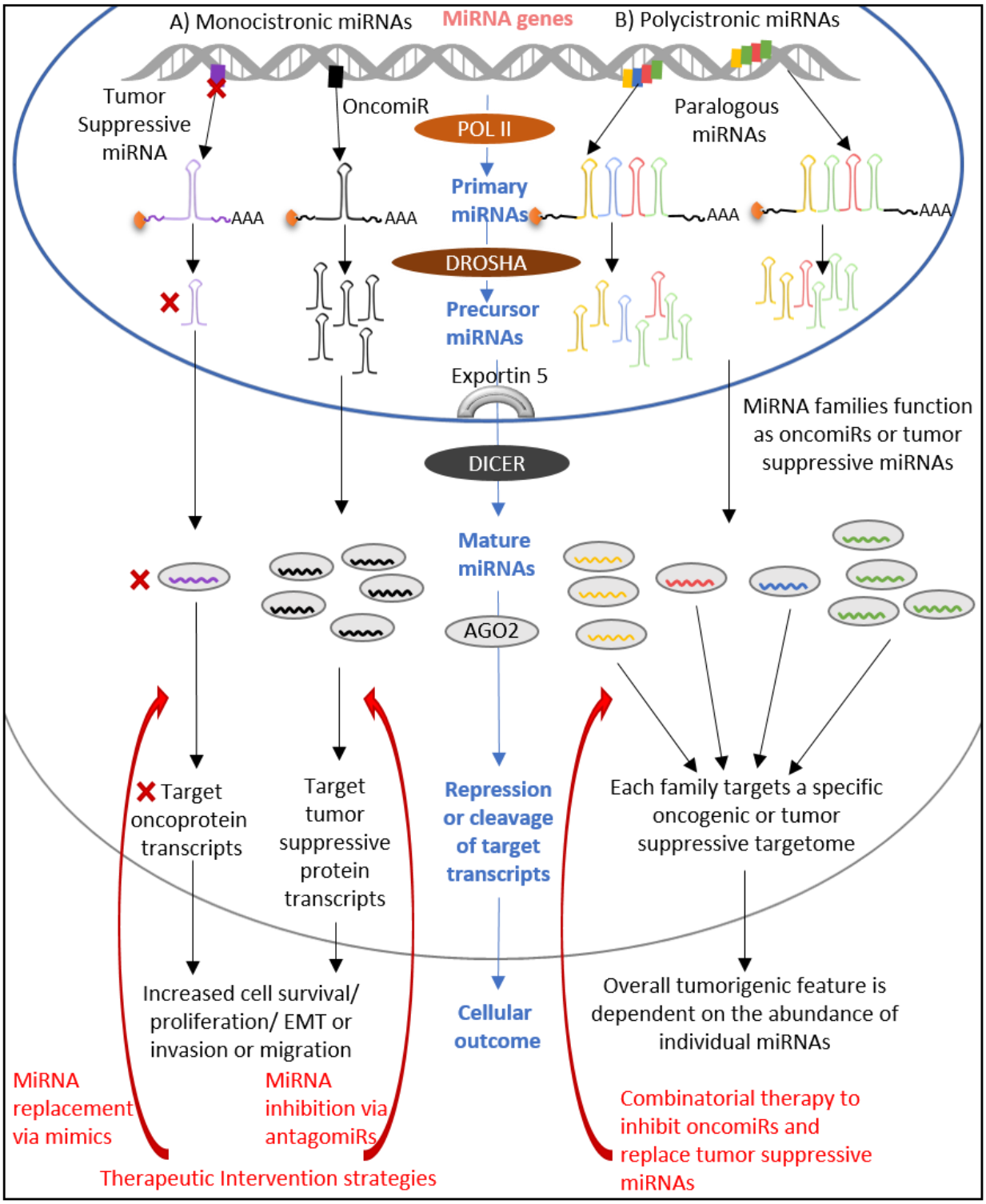
chromosomal region containing *mir-21* gene, a commonly overexpressed miRNA in multiple cancers(46) is an amplified CAGR (Reviews(23,24)).

Process of Biogenesis:

The primary miRNA (pri-miRNA) transcript produced as a result of RNA Polymerase II/III dependent transcription containing a single miRNA or as a cluster of miRNAs, produces a monocistronic or polycistronic pri-miRNA transcript, respectively(54). Pri-miRNA transcripts form stem-loop structures flanked by single-stranded (ss) RNA ends. For RNA Polymearse II transcripts, the ends contain a canonical 5′ 7-methylguanosine cap and a polyadenylation signaling at the 3′- end. The size of a typical pri-miRNA can range from a hundred to a few kilobases in length and can originate from either intragenic or intergenic miRNA genes(54,55) (**Figure 3.1**).

Processing of most pri-miRNAs begins with the association of the RNaseIII enzyme DROSHA and its cofactor Di George Syndrome Critical Region 8 (DGCR8) forming the microprocessor complex(56,57). The microprocessor complex recognizes the ssRNA regions of the pri-miRNA sequence flanking the stem-loop and cleaves the ends. The resultant ~60-80 nucleotide long hairpin structure is referred to as a precursor miRNA (pre-miRNA)(56,57). The pre-miRNA is translocated into the cytoplasm via Exportin-5 where another RNase III enzyme, DICER1 performs additional processing(58,59). DICER1 cleaves the pre-miRNA to generate a ~22 nucleotide duplex molecule containing the guide and the passenger miRNA strands. Following cleavage, the DICER1-miRNA duplex associates with Transactivation-Responsive RNA-binding protein (TRBP) that mediates a stable transfer of the miRNA duplex into an Argonaute protein (AGO) (60,61). Selective incorporation of the miRNA duplex into either AGO1, AGO2, AGO3 or AGO4 is dictated by the presence of bulges or mismatches in nucleotides 9-12 of the duplex(62,63). Incorporation of a miRNA duplex in AGO2, an AGO protein with endonuclease activity, results in selective cleavage of the passenger strand. The ssRNA guide strand is retained, and with AGO forms the mature miRNA-induced silencing complex (miRISC)(64,65). On the other hand, the endonuclease activity deficient AGOs, AGO1, AGO3 and AGO4, generate a functional miRISC by binding to the guide strand and separating the passenger strand based on thermodynamic instability. The released passenger strand is shunted for further degradation(66) (Biogenesis reviews(20,22,55), Argonautes reviews(63,67)).

Figure 3.1. Overview of oncomiRs or tumor suppressive miRNAs encoded as monocistronic or polycistronic genes, their involvement in tumorigenesis, and potential use as miRNA-based cancer therapeutics.(A) A monocistronic miRNA gene encodes a transcript containing a single primary miRNA. In cancers, one mechanism to alter the abundance of a mature miRNAs is through changes in transcription of the primary miRNA, where the expression of a tumor suppressive miRNA is downregulated, while that of an oncomiR is enhanced. A tumor-suppressive miRNA typically targets transcripts encoding oncogenic proteins, therefore miRNA replacement therapies using tumor suppressive miRNA mimics are currently being tested. OncomiRs on the other hand target tumor suppressor protein transcripts, and hence their inhibition via antagomiRs is also a potential miRNA-based therapeutic strategy. (B) Transcription of a polycistronic miRNA gene or a miRNA cluster results in a primary miRNA transcript containing multiple miRNAs. The duplication of a cluster, and expression of a more or less intact cluster from multiple genomic loci generates paralogous miRNAs. The resultant miRNAs from paralogues can be predominantly tumor-suppressive or oncogenic; however, their function is often largely context dependent – i.e. tissue-specific, temporally regulated, etc. The potential therapeutic strategy targeting miRNAs expressed from clusters depends on the abundance of the individual tumor suppressive or oncogenic miRNAs. Combinatorial miRNA therapeutics is a potential strategy currently being evaluated to combat tumorigenesis where the altered ratio between oncogenic/tumor suppressive miRNAs drives cancer development. (Adapted from (145)).


A) Monocistronic miRNAs
MiRNA genes
B) Polycistronic miRNAs
Tumor Suppressive miRNA
OncomiR
POL II
Paralogous miRNAs
Primary miRNAs
AAA
DROSHA
Precursor miRNAs
Exportin 5
DICER
MiRNA families function as oncomiRs or tumor suppressive miRNAs
Mature miRNAs
AGO2
Target oncoprotein transcripts
Target tumor suppressive protein transcripts
Repression or cleavage of target transcripts
Each family targets a specific oncogenic or tumor suppressive targetome
Increased cell survival/ proliferation/ EMT or invasion or migration
Cellular outcome
Overall tumorigenic feature is dependent on the abundance of individual miRNAs
MiRNA replacement via mimics
MiRNA inhibition via antagomiRs
Combinatorial therapy to inhibit oncomiRs and replace tumor suppressive miRNAs
Therapeutic Intervention strategies

3.1.2 Mechanisms of action and functions of miRNAs:

(a) Incorporation into miRISC and targeting*:*

The well-established role of functionally active miRISC is to negatively regulate transcription of the target protein-coding transcripts. The canonical mechanism by which miRISC performs its function depends on the extent of complementary binding between the 5'-end "seed region" of the miRNA, the 3'-UTR of the target mRNA transcript, and the enzymatic activity of the AGO protein(17,63,68,69). Perfect complementarity between the seed sequence, nucleotides 2-7 of the guide miRNA strand, and the target results in either degradation or translational repression of the target. The fate of the target transcript is dictated by whether the incorporated AGO displays catalytic activity or not, and whether additional complementarity occurs between the target and the miRNA. If a catalytically active AGO is incorporated into miRISC and the binding between the miRNA and the target are complementary between nucleotides 9-12 of the miRNA, then target cleavage will occur(65,70). Translational inhibition in the absence of target degradation occurs when the miRNA binds to its target via partial complementarity (Review(17,63,69)) or if an endonuclease deficient AGO is included in mRISC. The partial complementarity between the miRNA and its target is highly conserved across species, providing the basis for a combinatorial interactome. A combinatorial interactome is the mechanism by which a single miRNA regulates multiple targets, thus simultaneously exerting its regulatory effects on various signaling pathways. For example, the very well-studied miRNA miR-21 simultaneously targets transcripts of proteins that regulate cell division and apoptosis, such as phosphatase and tensin homolog (PTEN)(71), and programmed cell death 4 (PDCD4)(72) to drive the process of tumorigenesis (Review(47)). Partial complementarity between the miRNA and the target also facilitates targeting of a single transcript mRNA by multiple miRNAs resulting in enhanced repression of the target. This resulting moderate-to-severe downregulation of target transcripts via a miRISC is the canonical mode of action of miRNAs (Reviews(55,69,73)).

(b) The role of family members in expression and targeting:

The mechanism of action of miRNAs originating from a single locus, or a single mature miRNA originating from multiple loci remain largely unchanged. In these instances, the same cohort of target mRNA transcripts is repressed (47,71,74,75). However, miRNAs originating from several

loci that contain subtle variations in their mature sequences (55,76) can exert their repressive functions on a larger repertoire of target transcripts. These miRNA families have acquired an evolutionary advantage relative to miRNAs with a single mature miRNA sequence. In addition to an increased pool of potential targets, the presence of multiple miRNA family members across the genome may allow at least one of the family members to evade transcriptional or epigenomic regulation. Therefore, the presence of multiple genetically distinct miRNA family members may prevent the depletion of an entire pool of a specific mature miRNA during the onset of a diseased state. For example, transcription of the twelve *let-7* miRNA genes produces nine unique mature miRNA sequences that differ by at most three nucleotides (Reviews(76,77)). While these minor nucleic acid changes can potentially alter the targeting affinity of the various family members (**Figure 3.1** depicts miRNA family members) many of the targets are shared. Thus, reduced expression of a single family member is not expected to generate profound phenotypic consequences. For example, the promoter of *let-7a3* resides in a heavily methylated region of the genome in normal cells resulting in low levels of *let-7a3* in a normal cell. This is however not the case for the other *let-7* family members which are highly expressed under normal conditions leading to a stably differentiated state of the cell. Nonetheless, upon the onset of tumorigenesis, the methylation state of cells become severely disrupted, and except for *let-7a3* all the other *let-7* isoforms become repressed(24,43,76). Thus, the presence of multiple genetic loci encoding miRNA family members and slight variations in sequence between members adds an additional layer of complexity in the regulation of miRNAs in adverse cellular conditions.

Similar to the *let-7* family, another well-studied family of miRNAs is the miR-34 family. The three canonical miR-34 family members include miR-34a that arises from a monocistronic locus, and miR-34b/c, which are expressed from a polycistronic transcript(25,27,35) (**Figure 3.1** depicts mono- and polycistronic miRNA genes). The function of miR-34 in normal physiology is well established as an inducer of cellular senescence and cell cycle arrest(38) (Review(27)). Nevertheless, the advantage of multi-loci encoding miR-34 family members is that miR-34a and miR-34 b/c can be differentially regulated in tissue specific context(37,78,79).Recent reports suggest that the tissue-specific expression of the miR-34 paralogues miR-449a/b/c add an additional level of complexity to the control of cancer cell proliferation, invasion, and migration(80-82). Indeed, it was not until the paralogue *mir-449a/b/c* cluster was deleted in mouse models that the *mir-34a*, *mir-34b/c* double mutant displayed a phenotype(83-85).

(c) The role of miRNA clusters and paralogous in targeting:

Analogous to the overlapping role that miRNA family members have on gene expression, some paralogous clusters can also have overlapping roles while others have gained novel functions. A paralogous miRNA cluster is generated when a cluster undergoes duplication and translocates to another area of the genome (**Figure 3.1** depicts miRNA paralogues). The resultant paralogue may express miRNAs similar to the parent cluster, located in relatively analogous positions(86-88). One such miRNA cluster, miR-17~92 has been extensively studied due to its implication in the human developmental syndrome, Feingold disease. Loss of *mir-17~92* results in severe skeletal abnormalities, and learning and developmental disabilities associated with Feingold disease(89). However, similar developmental defects were not observed following the knockout of two *mir-17~92* paralogous clusters– *mir-106b~25* and *mir-106a~363*(90). Additionally, the presence of a single wild-type *mir-17~92* allele was capable of mitigating the deleterious effects of the loss of *mir-17~92*, despite the absence of its paralogues(91). Collectively the three paralogous clusters encode a total of fifteen miRNAs that can be sub-classified into four miRNA families that are presumed to target analogous target transcripts. However in this case it can be inferred that alterations in certain nucleotides of the paralogues may have ceased their ability to compensate for *mir-17~92* deletion(87,90,91) (**Figure 3.1** depicts miRNA paralogues). Therefore, in order to dissect the function of each miRNA in a family of miRNAs or within paralogues demands the generation of appropriate model systems to advance the field forward.

3.1.3 MiRNA function and relevance in cancer

MiRNAs are important players in the normal developmental processes of animal species. As such, disruption in the normal physiological levels of certain miRNAs can lead to the development of multiple diseases, including cancers.

Detailed characterization of various miRNAs has revealed many important properties of these powerful post-transcriptional modulators in both normal and diseased states. In the context of cancers, certain miRNAs have been identified as functional “drivers of cancer”, whereas others are regarded as mere “passengers” in the tumorigenic process. A few known miRNA drivers of cancer become upregulated while others are severely downregulated or lost. The miRNAs that

promote hallmarks of cancer are referred to as oncogenic miRNAs (oncomiRs). Those that prevent or reduce tumorigenesis are collectively called tumor suppressive miRNAs (**Figure 3.1**).

OncomiRs

OncomiR coding genes are frequently located in regions of the genome that are aberrantly amplified, or are subject to increased expression(92). Increased expression of an oncomiR can be attributed to enhanced transcription of the oncomiR gene due to (i) availability of transcription factors, (ii) hypomethylation of its promoter, or (iii) its location in an intra- or intergenic region that is subject to increased expression in cancer via other mechanisms. OncomiRs can also be upregulated due to defects in biogenesis and/or stability of the mature miRNA(22,23,55,73,91). The way by which an oncomiR typically functions is through targeting tumor suppressive protein-coding transcripts via canonical mechanisms, or through other less understood non-canonical mechanisms.

The first oncomiR to be validated was the miR-17~92 cluster (oncomiR-1). Overexpression of the cluster led to the development of lymphoproliferative and auto-immune diseases in mice via targeting of BIM, a pro-apoptotic protein(93). Other targets of miR-17~92 that support the oncogenic role for this cluster include *PTEN*, *E2Fs*, and *MYC*. More detailed analysis of this cluster confirmed that the cell-type and context specific processing of individual miRNAs from the cluster adds an additional level of complexity to the function of the oncomiR(87,94). Co-operatively the individual miRNAs processed from miR-17~92 functions as an oncomiR. However, miR-92 alone can antagonize an additional cluster member, miR-19 and also negatively regulates the oncogenic effects of c-Myc(91,94,95). Because miR-19 alone can recapitulate the oncogenic role of the intact mir-17~92 cluster(95,96), negative regulation by miR-92 suggest that miR-92 may be functioning as a tumor suppressor. The function of the miR-17~92 miRNA cluster is extremely intriguing and is currently under active investigation. Specifically, molecular roles and tissue specific effects of individual miRNAs of the miR-17~92 cluster are being determined in appropriate model systems (90,91,95-98). More recently, the unprocessed, so-called intermediate forms of the cluster that are thought to be responsible for sequestering some of the miRNA biogenesis machinery away from other miRNAs, resulting in the well-reported global down-regulation of miRNAs observed during tumorigenesis was reported ((99) and unpublished work). These positive findings highlight the

importance of carefully dissecting individual miRNAs from clusters so as to accurately identify the functions of each of the miRNAs contained within them.

Other miRNAs that have been well established as oncomiRs due to their implication in multiple solid tumors and hematological malignancies are miR-21 and miR-155(100-102). Independent studies determined that overexpression of individual miRNAs such as miR-21 and miR-155 are sufficient to cause lymphoproliferative diseases. The mechanism by which miR-155 initiates cancer is not well understood, however, in leukemic mouse models it was determined that miR-155 promotes cancer progression, perhaps through gradual downregulation of its targets, SHIP and C/EBP(103). In miR-21-dependent mouse models of lung cancer or pre-B-lymphoma, downregulation of the miR-21 targets PTEN and PDCD4 (negative regulators of cell death and cell-cycle, respectively) contributed to enhanced proliferation and growth(47,71,72).

Tumor Suppressive miRNAs

About 50% of the miRNAs involved in repressing oncogenic protein-coding genes are located in, or are close to, fragile regions of the genome that are frequently deleted in cancer. Additional mechanisms elicited by cancer cells to repress tumor suppressive miRNAs include LOH, hypermethylation of the promoter, or the activation of transcriptional repressors that specifically downregulate the expression of the miRNA gene(23,25,38,52). In the case of most tumor suppressive miRNA genes, identification of their role in development preceded their characterization as tumor suppressors. The most striking example of such a tumor suppressive miRNA is *let-7. Let-7* was identified as a crucial differentiation factor in *C. elegans* prior to its identified role in tumorigenesis. Indeed, the development of cancer requires a reversal of a well-differentiated state to an undifferentiated state, thus, it is perhaps not surprising that downregulation or loss of *let-7* family members is common in tumorigenesis(76,77).

High levels of *let-7* expressed from multiple genomic loci are expected in normal fully differentiated cells(76,77). This results in repression of *let-7* targets which are important oncogenes, such as *KRAS*, *NRAS*(104,105), *HMGA2*(105,106), *LIN28*(107-110), and *MYC*(111). A candid tumor-suppressive miRNA, such as *let-7* has multiple loci of origin in order to maintain an appropriate level of the tumor suppressive miRNA as a defense mechanism against developing cancers(76,112). However, since most *let-7* isoforms are located in regions of the genome

frequently deleted in cancer, *let-7* is severely downregulated in multiple cancers(52). One anomaly to this rule is the expression of *let-7a-3* gene. In lung adenocarcinoma, epigenetic regulation of the gene encoding *let-7-a3* results in hypomethylation of the promoter enhancing the accumulation of the *pre-let-7a-3* transcript in lung cancer cells, and subsequently its potential oncogenic effects(24,43).

Additional miRNAs that have been well established as tumor suppressors include miR-15a and miR-16-1, which were among the first miRNAs that demonstrated a negative correlation with the development of cancers(52,113). *Mir-15a~16-1* is located in 13q14.3, a region that is homozygously or hemizygously lost in over 50% of CLL cases(113). Early reports correlated loss of *mir-15a~16-1* with an increase in expression of *BCL2*, a pro-survival factor that normally prevents cell death(114-117). In addition to targeting *BCL-2*, targets of miR-15a~16-1 include several cell-regulatory proteins, such as MCL1(117,118), another BCL2-family member, CCND1(116), a cell cycle regulator, and WNT3A(116), a protein that induces several tumorigenic features including survival, proliferation, and invasion. Hence it is speculated that the simultaneous overexpression of these pro-survival onco-proteins, as a result of the loss of *mir-15a~16-1* cluster may synergistically contribute to the development of cancers(113,119). To further evaluate the tumor-suppressive potential of each miRNA in the *miR-15a~16-1* cluster, modelling CLL in more sophisticated model systems is required.

3.1.4 MicroRNAs in the development of non-small cell lung cancer (NSCLC)

Several genes are reported to be causal of lung cancer, such genes are regarded as “drivers” (120,121). Proto-oncogenes are inactive oncogenes performing their normal cellular growth and survival functions, however, upon activation function as oncogenes and drive the process of cancer(122). On the other hand, genes that are negative regulators of cancer normally functioning by activating cell death and apoptotic pathways upon receiving death signals, tumor suppressor genes are frequently lost in cancer (123). Multiple mechanisms have been reported to be associated with activation of oncogenes. For example hepatocyte growth factor receptor (MET) activation has been shown to depend on ligand-mediated activation (124), somatic mutation causing constitutive activation of MET (125), or activation of the gene locus or the proto-oncogene, MET (122). In the case of a tumor suppressor such as PTEN a few of the multiple reported mechanisms

include somatic mutations in PTEN leading to loss of function(126), or promoter-hypermethylation resulting in abrogated transcription of PTEN(123). However, research from the last two decades have implicated the importance of miRNAs in positively regulating oncogenes, and negatively regulating tumor suppressor genes, as also described in the previous section.

The oncogene MET is regulated via various tumor suppressive miRNAs, one such example being miR-34 (127), that is frequently lost during the development of lung cancer (79,127,128). Moreover, in order to prevent cancer progression, some tumor-suppressive miRNAs have been reported to function in synergy. For example, miR-34 and *let-7* (127) or miR-34 and miR-15a~16-1 cluster (128) together efficiently negatively regulate oncogene activation or cell cycle progression, thereby suppressing lung cancer development. To add to the versatile mechanism of action of tumor-suppressive miRNAs in lung cancer, owing to a host of target recognition by a single miRNA, it has become apparent that miR-34 can negatively regulate multiple oncogenes (MET and Epidermal Growth Factor Receptor (EGFR)), simultaneously to prevent lung tumorigenesis (129).

On the other hand, oncomiRs that are reported to repress lung cancer tumor suppressor genes, are frequently reduced in expression during the process of lung cancer development and progression. One such *bona fide* oncomiR, miR-21 that is frequently overexpressed in lung cancer, negatively regulates the tumor suppressor gene, *PTEN* (71). Apart from *PTEN*, miR-21 is also reported to target additional tumor suppressors such as programmed cell death 4 (PDCD4), thereby promoting lung cancer development via a concerted effect (72,130). Additionally, multiple other oncomiRs such as miR-17~92cluster, miR-155, miR-221/22, along with a few others have also been described to promote lung tumorigenesis (131-133).

3.1.5 MicroRNAs in TKI resistance in NSCLC

In the development of resistance to drugs utilized to treat lung cancer, multiple miRNAs have been implicated to have a role. A few miRNAs that are known to target transcripts of proteins associated with establishment of various hallmarks of cancer (**Figure 3.2**), have also been reported to cause resistance to erlotinib, a Tyrosine Kinase Inhibitor (TKI) utilized as standard-of-care for NSCLC patients with constitutively active Epidermal Growth Factor Receptor or other TKIs. For example, the oncomiR, miR-21 is severely upregulated in TKI-resistant NSCLC cells, resulting in

downregulation of PTEN and PDCD4, resulting in the enhancement of cell proliferation via activation of the pro-growth AKT pathway (134,135). In contrast, the *bona fide* tumor-suppressor miRNA, miR-34 is a miRNA that is recurrently downregulated in TKI resistant patients (136-138). Therefore, restoring miR-34 to re-sensitize TKI resistant NSCLC cells it is under active investigation, either alone (136) or in combination with erlotinib or other miRNAs that synergize with miR-34 to induce cell-cycle arrest and apoptosis of resistant cells (137,138).

Metabolic reprogramming of cells also mediates development of resistance to TKIs, by enhancing cell proliferation and survival pathways (139-141). Although a few miRNAs involved in NSCLC development can perform their role by altering metabolic phenotypes of cells, such as miR-124 (142), not many miRNAs have been linked to have a role in mediating resistance to cancer drugs through this mechanism. Two miRNAs thus far, miR-133b (143) and miR-513a-3p(144), are known to be severely downregulated in cisplatin treated lung cancer cells, and restoring the miRNAs individually can re-sensitize the cells. It was demonstrated that these miRNAs target glutathione S-transferase P1 (GSTP1), an enzyme actively involved in xenobiotic metabolism, thereby loss of the miRNAs enhanced tolerance of lung cancer cells to cisplatin (143,144).

Lastly, to the best of our knowledge, only one recent study thus far has identified the role of a miRNA in mediating TKI resistance via metabolic reprogramming of cells. This study by Slack and colleagues (2019) demonstrated that overexpression of miR-147b led to an altered metabolic landscape in cells that develop tolerance to a TKI. MiR-147b altered the tricarboxylic acid (TCA) cycle that favored cell survival by activating pseudohypoxic signaling due to increased accumulation of succinate in the TKI resistant population (139). Although only one miRNA has thus far been linked to TKI resistance mediating its effects via altering key metabolic pathways, it is not surprising that other miRNAs reported to function as *bona fide* modulators of TKI response may also function similarly. It can be hypothesized so because one miRNA can alter or affect multiple targets simultaneously, and by analogy, can regulate multiple signaling pathways at the same time.

Currently, over 2,656 mature miRNAs have been identified (miRbase, version 22)(146). MiRNAs that become dysfunctional leading to the development of cancer, regarded as oncomiRs

or tumor-suppressive miRNAs have been extensively studied, implicating their specific roles in development of various hallmarks of cancer. However, only a handful of miRNAs have thus far been reported to mediate the development of resistance to current therapies in lung cancer, such as erlotinib. Therefore, in this study, a miRNA screen is conducted in a TKI sensitive cell line to identify novel miRNAs that induce erlotinib resistance.

3.1.6 Study design and hypothesis

Erlotinib treated patients with constitutively activated EGFR develop resistance to treatment by multiple mechanisms (Chapter 1, **Figure 1.8**). Development of secondary mutations in EGFR and activation of alternative pathways bypassing EGFR signaling (described in detail in chapter 1) have been reported to be causal of acquired resistance to erlotinib in majority of the patients. However, in over 15-20% of cases the mechanism mediating erlotinib resistance remain unknown. Therefore, the hypothesis of this study is that novel epigenetic factors such as miRNAs can mediate erlotinib resistance upon overexpression. To this end, 2,019 individual miRNAs were overexpressed in an NSCLC cell line that is sensitive to TKIs, including erlotinib (identified through National Cancer Institute was used (NCI-60, DTP) (149)).The miRNAs that enhanced cell proliferation post-transfection were further validated. The predicted targets of the miRNA of interest were bioinformatically evaluated to predict relevant targets and pathways that mediate erlotinib resistance.

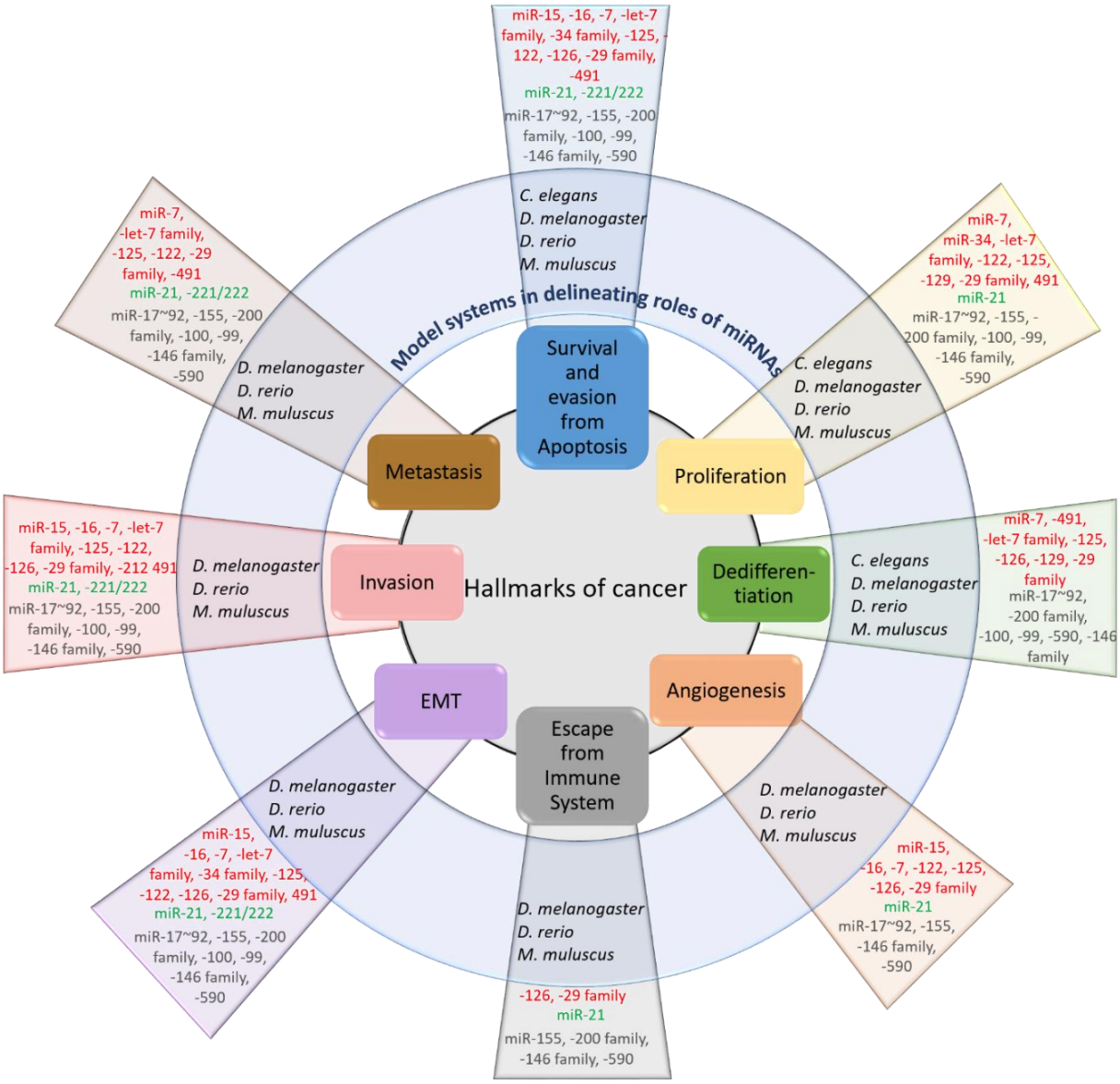


Figure 3.2. Functions of miRNAs in regulation of the hallmarks of cancer, identified via the use of *in vivo* model organisms. Hallmarks of cancer are the cellular processes that become severely dysregulated upon the onset of a cancer. The various model organisms, owing to their endogenous properties have been utilized to delineate the functions of the enlisted miRNAs that mediate the specific cancerous feature. MiRNAs represented in red are *bona fide* tumor suppressive miRNAs, in green are oncomiRs, while in grey represent miRNAs that, depending on their context, can function as either a tumor suppressive or oncogenic miRNA (Adapted from (145)).

3.2 Methods:

3.2.1 Cell culture

All cell lines utilized in the study were obtained from American Type Culture Collection (ATCC), cultured under standard conditions and were confirmed to be free of mycoplasma. Cell lines generated during the study were authenticated by ATCC Cell Line Authentication and grown in RPMI media supplemented with 10% FBS and 1% Penicillin/Streptomycin cocktail. Cell lines generated during this study, EKVX-pmiR and H322M-pmiR, were continuously cultured in media containing 16 or 8μg/ml G418 respectively for EKVX-pmiR and H322M-pmiR cells. However, the screen was conducted in phenol free RPMI media supplemented with 10% FBS and 1% Penicillin/Streptomycin.

3.2.2 Generation and characterization of cell lines

Erlotinib sensitive cell lines, EKVX and H322M were forward transfected with 2μg of linearized pmiRGLO plasmid (Promega) using lipofectamine 2000 (11-668-019, Thermo Fisher Scientific), as per manufacturer's instructions. Forty-eight hours later, cells were selected using 100μg/ml G418. EKVX-pmiR and H322M-pmiR cells stably expressing pmiRGLO plasmid were clonally selected. Ten-thousand cells for each single clone were plated in a 96-well plate in 6 replicates, and 32 hours post plating firefly and renilla activities were measured using Dual-GLO Luciferase assay kit (Promega, E2920) following manufacturer's protocol (**Figure 3.3 A, B**). Renilla activity in EKVX-pmiR Clone 2 and H322M-pmiR clone 1 (further regarded as EKVX-pmiR and H322M-pmiR, respectively) were further investigated as proxy for cell number linearity, by plating increasing cell numbers in a 384 well plate, using Dual-GLO Luciferase assay (**Figure 3.3 C**). Finally, 2,000 EKVX-pmiR or H322M-pmiR cells in 6 replicates in a 384 well plate were reverse co-transfected using 0.6 nM silencing RNA targeting luciferase reporter (siLUC2) or negative control (sicont) transfected with 6 nM premiR-control (negative control for miRNAs) using RNAiMAX, following the manufacturer's protocol, to evaluate transfection efficiency by measuring repression of firefly activity, using Dual-GLO Luciferase assay (**Figure 3.4 A**). Final characterization was performed by evaluating erlotinib response of clones relative to parents, described below in *erlotinib dose response* (**Figure 3.3 D**).

3.2.3 Drug Preparation for *in vitro* studies

Erlotinib (S7786, Selleck Chemicals), was dissolved in DMSO to prepare 0.4 M stock solutions, which were aliquoted and stored in -80°C. Working concentration of all the drugs was 200 µM prepared in complete medium and diluted to different concentrations for *in vitro* experiments.

3.2.4 Selection of controls for the overexpression screen

MiR-21 (mirVana miRNA mimic, Life Tech, Catalog # 4464066, Assay ID # MC10206) or miR-17 antagomir (Anti-miR miRNA Inhibitor, Life Tech, Catalog # AM12412, Assay ID # AM17000), which are known mediators of erlotinib resistance were inconsistent at inducing erlotinib resistance in EKVX-pmiR cells between experiments, and therefore we conducted the following experiment to select appropriate positive controls for the screen. The human mirVana library of miRNAs (Invitrogen; based on miRBase v.21) contains 2,019 miRNAs individually arrayed in 96 well plates. One of 23 plates that make up the library was taken at random (Hs Mimic v19-A4-4) and the below described screening procedure was conducted using four biological replicates. Two miRNAs (miR-219b-3p and miR-4749-5p) that enhanced cell growth relative to an arbitrary value assigned to the negative control (Mean+ 4X standard deviations) in all four replicates were considered as good-fit for positive controls to be used in the screen (**Figure 3.5**).

3.2.5 MiRNA overexpression screen

EKVX-pmiR cells (2 x10^3) were reverse co-transfected with 6nM premiR-control (negative control) or miR-219b-3p and miR-4749-5p (positive controls) or the miRNAs from the human mirVana library along with 0.6nM siLUC2 in a 384-well plate using 0.1ul RNAiMAX in a final volume of 10µL media, in 6 replicates (based on results from **Figure 3.4 B**). Transfection with siLUC2 was used to evaluate transfection efficiency and for normalization. Each plate included the positive and negative controls. Twenty-four hours post transfection, media containing a final concentration of 75% growth inhibitory (GI75) concentration of erlotinib or equivalent DMSO (negative control) was added to the positive and negative control transfected wells to validate efficacy of the drug and the pro-growth effect of positive control miRNAs. Simultaneously, GI75 erlotinib containing media was added to the miRNA transfected wells. Seventy-two hours post-treatment, firefly and renilla activities were measured using Dual-GLO Luciferase assay kit

(Promega, E2920). Transfection efficiency for each transfection was calculated as described in **Figure 3.3 D**. Following which, renilla activity of transfected cells was calculated. The growth of erlotinib treated miRNA transfected cells is represented relative to that of negative control transfected cells. Results from the 23 individual mirVana plates were compiled, and miRNAs that enhanced cell growth relative to an arbitrary value assigned to that of the positive controls i.e. Mean plus 4X standard deviations were further re-evaluated (**Table 3.1, Figure 3.6**).

3.2.6 Validation of candidates by overexpression and knockdown experiments

To evaluate the effect of miR-432-5p in mediating erlotinib resistance in sensitive NSCLC cells, 6nM miR-432-5p mimic (mirVana miRNA mimic, Life Tech, Catalog# 4464066, Assay ID: MC10941) was reverse transfected using RNAiMAX. To evaluate the effect of miR-432-5p in re-sensitizing resistant NSCLC cells to erlotinib, 150nM hsa-miR-432-5p antagomir (Anti-miR miRNA Inhibitor, Life Tech, Catalog # AM17000, Assay ID # AM10941) was reverse transfected using RNAiMAX.

3.2.7 Bioinformatic analysis

Ingenuity pathway analysis (IPA), v01-12 was performed using experimentally validated targets of the top 60 validated miRNAs **(Table 3.1),** or for miR-432-5p. Lung cancer Illumina hiseq (miRNA seq) data was retrieved from TCGA (LUAD and LUSC) using R x64 v3.3.3. and plotted using GraphPad Prism version 6 software (GraphPad Software). MiRNA target prediction softwares, TargetScan Human (147) and mirmap (148) were utilized to identify miR-432-5p putative targets. NF1 alteration frequency in LUAD (TCGA) was evaluated using cBioPortal (154).

3.2.8 RNA isolation and Quantitative real time PCR (qRT-PCR)

Total RNA was isolated from $1.5X10^5$ cells in 6 well plates or $1X10^6$ cells grown in 10-cm plates after 72 hours post-transfection using the miRneasy Kit with DNase I digestion. RNA integrity was evaluated on a 1.5% agarose gel, and total RNA quantified using a nanodrop. cDNA was then synthesized from 1μg of total RNA isolated from cells using miScript Reverse Transcriptase kit, as indicated by the manufacturer's protocol (Qiagen). MiScript SYBR Green PCR Kit (Qiagen)

was used as indicated by the manufacturer's protocol, to quantify target miRNA expression normalized to the housekeeping, RNU6B via qRT-PCR. RNU6B (control, MS00033740, Qiagen) and miR-432-5p (MS00031850, Qiagen)

3.2.9 Erlotinib dose response

The protocol followed to evaluate Erlotinib dose response of various NSCLC cells and cells generated in this study was as per the NCI-60 Cell Five-Dose Screen (NCI-60, DTP) (149). Briefly, Sulforhodamine B colorimetric assay (SRB assay) was performed by exposing cells to varying concentrations of Erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. Post data normalization, as described in figure legends, GI50 erlotinib was calculated from the respective dose curves (NCI-60, DTP) (149).

3.2.10 Statistical analysis

All data were analyzed using GraphPad Prism version 7 software (GraphPad Software) and are presented as mean values ± SEM. Student's t-test or one-way ANOVA were the statistical analyses performed, as specified in the figure legends. P-value of < 0.05 was considered significant.

3.3 Results:

3.3.1 Cell line generation to conduct overexpression screen and validation experiments

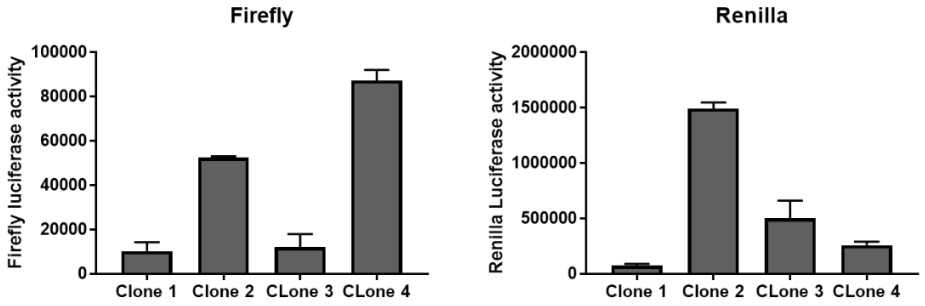
Two cell lines, EKVX and H322M, reported to be sensitive to erlotinib by the study conducted by NCI-DTP (149) and validated in our lab, were utilized in this study. To monitor the growth and transfection efficiency of the cells, EKVX and H322M cells were generated that stably expressed pmiRGLO (Promega), a plasmid that co-expresses both renilla and firefly luciferase. Renilla was used as a proxy for cell number while firefly was a marker of transfection efficiency. Cells selected via clonal selection were evaluated for their luciferase activities, described in **Figures 3.3 A and B**. EKVX-pmiR clone 2 and H322M-pmiR clone 1 with high levels of both firefly and luciferase activities were further characterized (further regarded as EKVX-pmiR and H322M-pmiR, respectively). To evaluate if renilla can be utilized to represent cell number in the reporter system, pmiRGLO expressing cells were plated at increasing numbers and renilla activity was measured

(**Figure 3.3 C**). EKVX-pmiR and H322M-pmiR cells were also evaluated for their response to erlotinib that closely recapitulated that of the parental EKVX and H322M cells (**Figure 3.3 D**). EKVX-pmiR cells were used to perform the overexpression screen, whereas both EKVX-pmiR and H322M-pmiR cells were utilized for validation of candidate miRNAs.

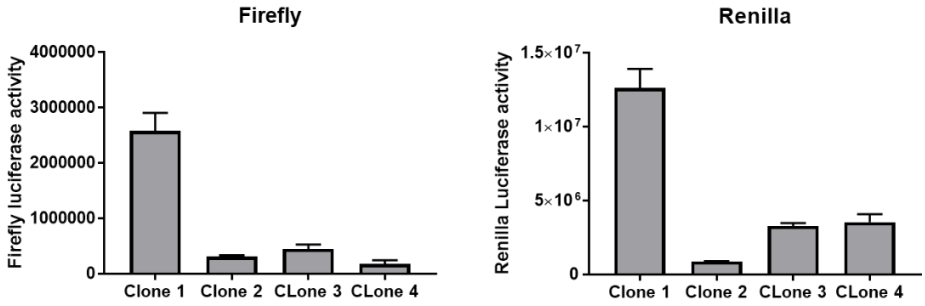
To determine if targeting firefly with an siRNA would serve as an appropriate normalizer for well-to-well variability in transfection, 10% of the transfection cocktail was supplemented with siLUC2. For example, 6nM of premiRcontrol (miRNA negative control) along with 0.6nM of silencing RNA to the luciferase reporter gene (siLUC2) or a negative control siRNA (sicont) were co-transfected and luciferase reporter activity was measured and calculated relative to renilla activity (cell number). Average transfection efficiencies for EKVX-pmiR=77.8% and H233M-pmiR=71.02%, **Figure 3.4 A**. Ratios of premiRcontrol to siLUC2 with 0.05 or 0.1 µl of RNAiMAX were also conducted to determine the ratio at which maximum efficiency was achieved. Based on this analysis, 6nM premiRcontrol co-transfected with 0.6nM siLUC2 using 0.1ul RNAiMAX was further utilized to conduct the miRNA overexpression screen (**Figure 3.4 B**).

Figure 3.3. Clonal selection and characterization of EKVX and H322M cells stably expressing the reporter pmiRGLO plasmid. A) Clonally selected EKVX cells stably expressing pmiRGLO plasmid were evaluated for their respective firefly and renilla activities. B) Clonally selected H322M cells stably expressing pmiRGLO plasmid were evaluated for their respective firefly and renilla activities. C) Cell number linearity of EKVX-pmiR clone 2 and H322M-pmiR clone 1 cells was measured eighteen hours post seeding by evaluating renilla activity. D) Erlotinib dose response via SRB assay was evaluated by exposing the parental EKVX or H322M cells or the EKVX-pmiR clone 2 and H322M-pmiR clone 1 cells to varying concentrations of Erlotinib or the highest equivalent volume of dimethyl sulfoxide (DMSO, negative control) containing media for 72 hours. For percent of cells calculation, number of cells at the time of addition of Erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. Fifty percent growth inhibitory concentration of Erlotinib (GI50) was calculated from the respective dose curves (as per NCI-60 Cell Five-Dose Screen, NCI-60, DTP (149).

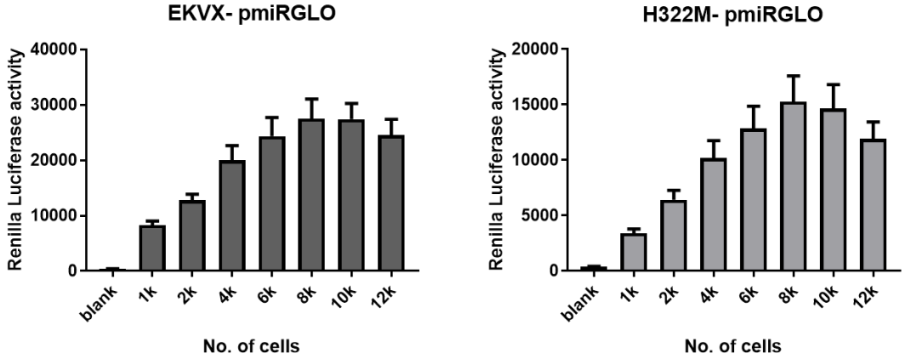
A.



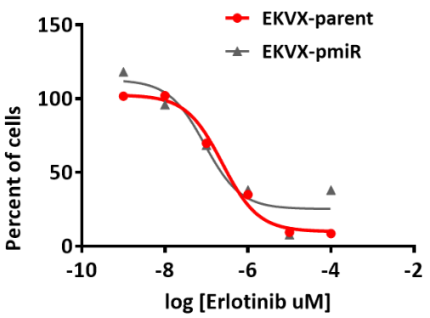
B.



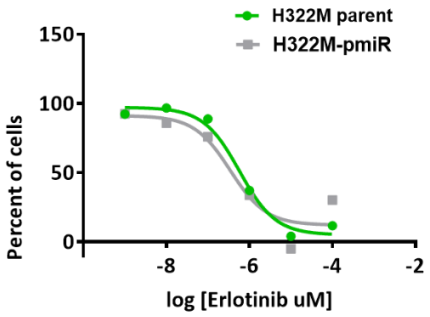
C.



D.



	parent	EKVX-pmiR
GI50	0.24 µM	0.91µM



	parent	H322M-pmiR
GI50	0.8 µM	0.3µM

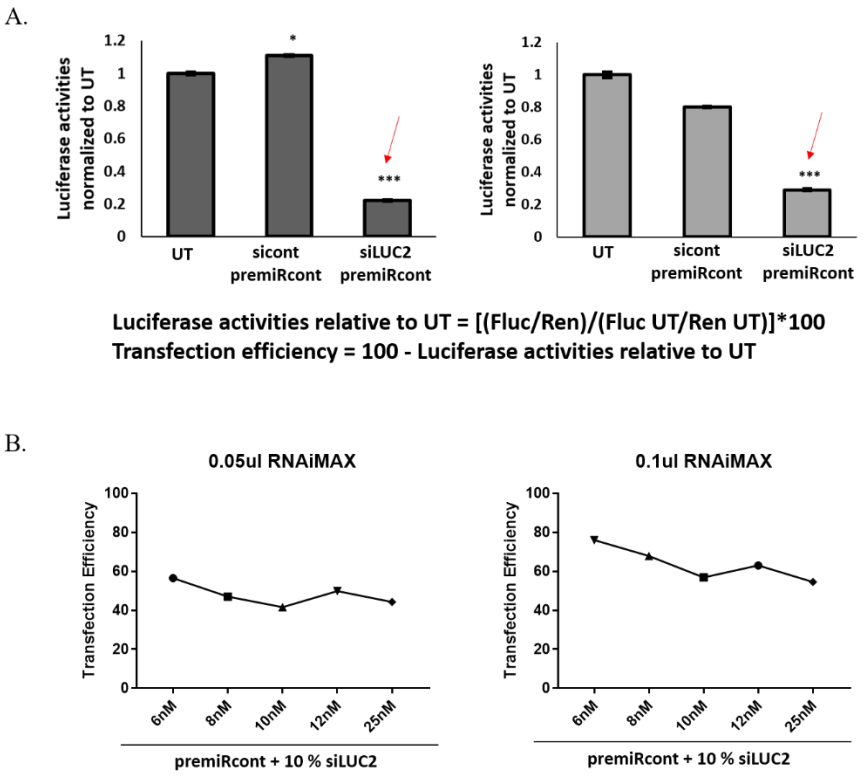


Figure 3.4. Determination of transfection efficiency of EKVX-pmiR and H322M-pmiR cells. A) EKVX-pmiR or H322M-pmiR cells were co-transfected with 6nM premiR control (premiRcont) and 10% siRNA control (sicont) or siRNA to Firefly Luciferase (siLUC2) using 0.1µL of RNAiMAX. Thirty-two hours post transfection, reporter activity was measured. UT= untransfected. Red arrows indicate calculated transfection efficiencies as per the inset equation. B) EKVX-pmiR cells were transfected using variable premiRcont to siLUC2 ratio using either 0.05 or 0.1 µL RNAiMAX.

3.3.2 Identification of positive controls

MiR-21 is reported to be a miRNA that causes acquired resistance to erlotinib in NSCLC (134,135). On the other hand, loss of miR-17 is known to mediate erlotinib resistance (150). However, in this study, by either overexpressing miR-21 or downregulating miR-17, a consistent pro-growth effect on EKVX-pmiR cells post erlotinib treatment was not observed. Therefore, a pilot screen was conducted in replicates of four, to identify at least two miRNAs that consistently enhanced cell growth of EKVX-pmiR cells, post erlotinib treatment (**Figure 3.5**).

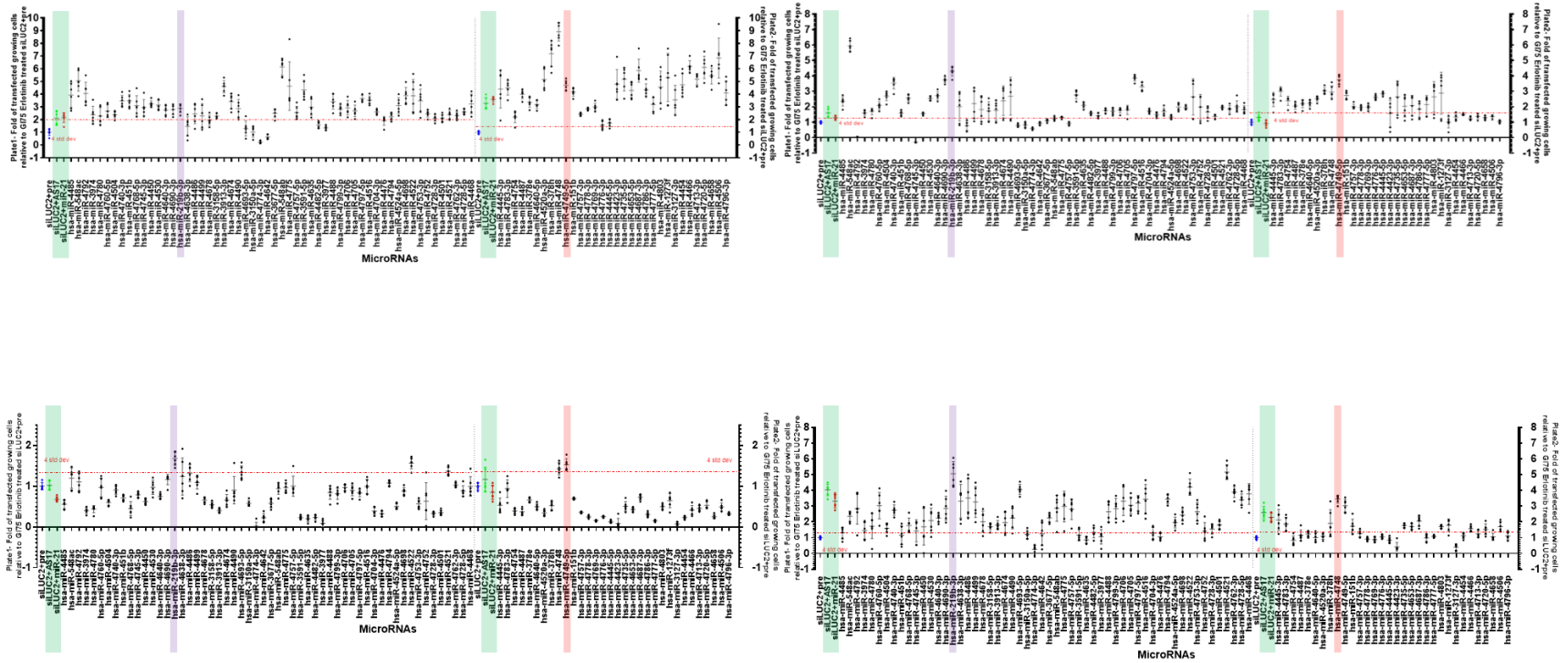


Figure 3.5. Selection of positive controls for miRNA overexpression screen. EKVX-pmiR cells were transfected with 6nM premiRcontrol (pre) or miR-21or anti-sense-17 (AS17) or miRNAs from mirVana library (Hs Mimic v19-A4-4). At the same time, all wells were transfected with 0.6nM siLUC2 for transfection normalization. Twenty-four hours post-transfection, cells were exposed to GI75 erlotinib concentration. Seventy-two hours post treatment, reporter activities were measured. Following normalization for transfection (See Figure 3.4A),the growth of erlotinib treated miRNA transfected cells is represented relative to that of negative control transfected cells. Cut-off (dotted red line) is represented as mean plus 4X standard deviation of the value of negative control. Highlighted in green are transfected cells in the presence of erlotinib post miR-21 or AS17 transfection, in purple and red are putative positive controls, miR-219b-3p and miR-4749.

3.3.3 MiRNA overexpression screen results

Since miRNAs have been reported to function as mediators of erlotinib resistance, 2,019 miRNAs were evaluated individually to identify novel mediators of erlotinib resistance. EKVX-pmiR cells were seeded in 384-well format and were transfected with 6nM of each individual miRNA. Twenty-four hours later cells were exposed to GI50 erlotinib for 72 hours. Transfected cells that displayed enhanced cell growth in the presence of the drug relative to the cut-off (**Figure 3.6**), were taken forward for re-evaluation of their pro-growth response to erlotinib.

3.3.4 Re-evaluation of pro-growth effect of top 60 miRNAs post erlotinib treatment in sensitive NSCLC cells

The miRNAs that positively regulate cell growth in the presence of erlotinib post transfection in EKVX-pmiR cells were re-evaluated in a second erlotinib sensitive NSCLC reporter line, H322M-pmiR cells. Data from both cell lines is represented in **Table 3.1**. Fifteen of the top 60 miRNAs induced cell death in the H322M-pmiR cells (negative fold growth), suggesting a few miRNAs may have a cell-specific role in erlotinib response.

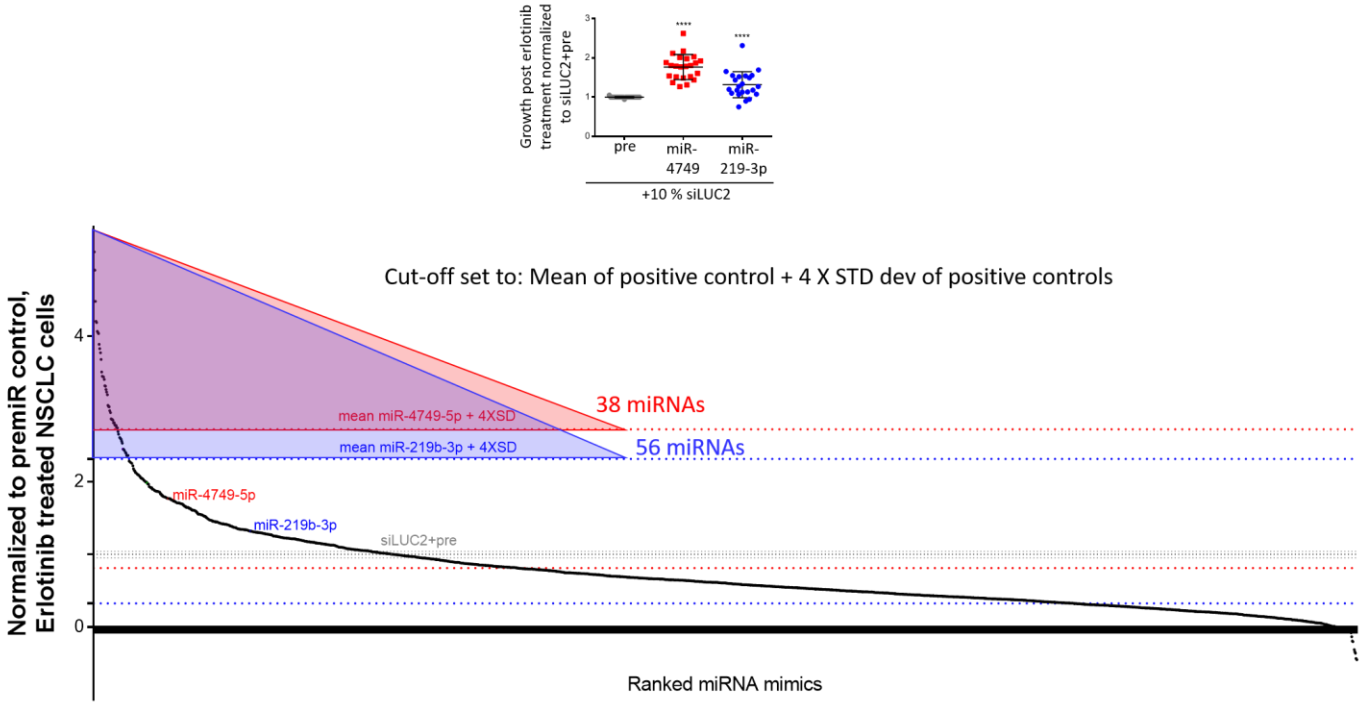


Figure 3.6. Overexpression of certain miRNAs enhance erlotinib resistance. EKVX-pmiR were co-transfected with 6nM premiRcontrol (pre) or miRNAs from mirVana or positive controls and 0.6nM siLUC2. Twenty-four hours post-transfection, cells were exposed to GI75 erlotinib concentration. Seventy-two hours post treatment, reporter activity was measured. Fold growth of transfected cells in the presence of erlotinib are represented relative to negative control (siLUC2+pre), in ranked order on the x-axis. Grey dotted lines represent mean + 4X standard of negative control, while red and blue dotted lines represent the same for miR-4749 and miR-219b-3p. Inset graph represents fold growth of premiRcontrol (pre) or positive controls (miR-4749 and miR-219b-3p) transfected cells in the presence of erlotinib from 23 plates of miRVana library used for the screen.

Table 3.1. Top 60 miRNAs that were evaluated for growth response in the presence of erlotinib, in EKVX-pmiR and H322M-pmiR cells. MiRNAs in bold resulted in heightened resistance to erlotinib in both cell lines, and constitute the top-five miRNAs selected for further validation.

EKVX-pmiR cells		**H322M-pmiR cells**	
Top 60 miRNAs	**Fold growth of transfected cells**	**Top 60 miRNAs**	**Fold growth of transfected cells**
hsa-miR-5693	**4.806839**	**hsa-miR-5693**	**6.980146**
hsa-miR-3618	**4.128858**	**hsa-miR-432-5p**	**5.680975**
hsa-miR-3657	3.53778	**hsa-miR-3618**	**5.201033**
hsa-miR-432-5p	**3.447473**	hsa-miR-4787-5p	3.098385
hsa-miR-4435	**3.369496**	hsa-miR-3198	2.666229
hsa-miR-588	**3.115648**	**hsa-miR-4435**	**2.599998**
hsa-miR-4701-3p	3.096875	**hsa-miR-588**	**2.331268**
hsa-miR-4787-5p	2.550594	hsa-miR-4693-5p	2.305733
hsa-miR-4521	2.406701	hsa-miR-4797-5p	2.231674
hsa-miR-4499	2.176414	hsa-miR-4775	2.208845
hsa-miR-9-5p	2.139334	hsa-miR-4521	2.115477
hsa-miR-4643	2.074379	hsa-miR-4643	2.072277
hsa-miR-4486	1.758585	hsa-miR-4781-3p	1.938728
hsa-miR-548ac	1.743734	hsa-miR-4757-5p	1.797722
hsa-miR-4757-5p	1.71694	hsa-miR-4328	1.77053
hsa-miR-3926	1.614334	hsa-miR-4638-3p	1.730049
hsa-miR-4760-5p	1.552469	hsa-miR-4329	1.706815
hsa-miR-4748	1.508298	hsa-miR-4634	1.588003
hsa-miR-4792	1.494049	hsa-miR-4499	1.475899
hsa-miR-3198	1.474083	hsa-miR-4791	1.467733
hsa-miR-5694	1.472514	hsa-miR-4792	1.452575
hsa-miR-4778-5p	1.463761	hsa-miR-3665	1.405999
hsa-miR-4638-3p	1.358205	hsa-miR-4701-3p	1.387164
hsa-miR-3677-5p	1.339528	hsa-miR-4753-3p	1.364724
hsa-miR-4794	1.313135	hsa-miR-4793-5p	1.352131
hsa-miR-4693-5p	1.313076	hsa-miR-4706	1.338108
hsa-miR-4797-5p	1.298536	hsa-miR-4698	1.306059
hsa-miR-4328	1.276663	hsa-miR-4488	1.281768
hsa-miR-4753-3p	1.207611	hsa-miR-4778-5p	1.251342
hsa-miR-4781-3p	1.206249	hsa-miR-4740-3p	1.224541
hsa-miR-4488	1.182468	hsa-miR-4330	1.168158
hsa-miR-4740-3p	1.181966	hsa-miR-4748	1.116061
hsa-miR-548ab	1.140096	hsa-miR-3677-5p	1.110286
hsa-miR-4793-5p	1.126522	hsa-miR-4794	1.103088
hsa-miR-4634	1.121035	hsa-miR-4705	1.09914
hsa-miR-4705	1.119791	hsa-miR-4760-5p	1.035345

Table 3.1. continued

hsa-miR-4791	1.107927	hsa-miR-4728-5p	0.980355
hsa-miR-4706	1.104596	hsa-miR-548ak	0.893768
hsa-miR-4522	1.043842	hsa-miR-3926	0.810352
hsa-miR-4775	1.017418	hsa-miR-5694	0.804238
hsa-miR-548ak	0.971604	hsa-miR-9-5p	0.732981
hsa-miR-4468	0.943862	hsa-miR-548ac	0.313635
hsa-miR-3665	0.934691	hsa-miR-548ab	0.232357
hsa-miR-548aj-3p	0.926283	hsa-miR-4676-3p	0.222049
hsa-miR-4751	0.891378	hsa-miR-4654	0.120642
hsa-miR-4330	0.883232	hsa-miR-4661-3p	-0.01684
hsa-miR-4654	0.851533	hsa-miR-4741	-0.02058
hsa-miR-4661-3p	0.844637	hsa-miR-4516	-0.18871
hsa-miR-4329	0.841927	hsa-miR-4751	-0.29721
hsa-miR-4698	0.837794	hsa-miR-204-3p	-0.30431
hsa-miR-4799-3p	0.821699	hsa-miR-4690-3p	-0.31774
hsa-miR-4728-5p	0.798065	hsa-miR-4486	-0.47874
hsa-miR-4534	0.788234	hsa-miR-548ad	-0.57008
hsa-miR-548ad	0.7357	hsa-miR-4468	-0.59024
hsa-miR-4762-3p	0.661747	hsa-miR-4522	-0.62091
hsa-miR-4516	0.544357	hsa-miR-4762-3p	-0.81152
hsa-miR-4741	0.529365	hsa-miR-3657	-1.16928
hsa-miR-204-3p	0.424363	hsa-miR-548aj-3p	-1.52673
hsa-miR-4690-3p	0.383958	hsa-miR-4799-3p	-1.57215
hsa-miR-4676-3p	0.331065	hsa-miR-4534	-2.21756

3.3.5 Bioinformatic and functional analysis of key pathways regulated by the top 60 miRNAs

Ingenuity pathway analysis (IPA), a software to analyze interactions and pathways that can be mediated by genes and miRNAs was performed for the top 60 miRNAs. The IPA function used for this analysis was "experimentally validated mRNA targets of miRNAs". Multiple cancer related pathways are predicted to be regulated by the targets of the identified miRNAs. In the predicted dysregulated networks, cell death and survival, cancer, and cell cycle were the highest scoring cell functions (**Figure 3.7 A**). A high scoring pathway is less likely to be identified at random. In this case, the highest score 39 suggests that the probability of obtaining this exact pathway with the certain predicted targets of miRNAs has one in $1X10^{39}$ chances. Additionally, in toxic functions mediated by the genes regulated by miRNAs, cell death and survival, and cell

growth and proliferation were significantly enriched with the involvement of multiple experimentally predicted targets (21 molecules in each) (**Figure 3.7 B**).

Lastly, target genes of the top 60 miRNAs significantly overlapped with known regulators of specific pathways. The top predicted canonical pathway to be altered by these miRNAs was cancer drug resistance by efflux, with at least 25 of 49 known molecules in the pathway being significantly involved. Additionally, pathways that are reported to be involved in resistance to erlotinib specifically, such as PTEN and PI3K/Akt signaling (151,152) are also altered by the top 60 miRNAs (**Figure 3.7 B**). Overlapping these predicted canonical pathways identifies a cross talk between all the pathways. Such a cross talk can provide information on novel cellular pathways that have never been identified to be involved in mediating resistance to erlotinib (**Figure 3.7 D**). Of the top 60 miRNAs that were found to promote erlotinib resistance, 38 are predicted to be involved specifically in cancer drug resistance by efflux pathway including four out of the top five miRNAs, shown in bold (**Table 3.2**)

Table 3.2. IPA curated 38 of 60 miRNAs known to be involved in cancer drug resistance by efflux pathway. MiRNAs in bold represent four of the top five miRNAs involved in promoting erlotinib resistance from Table 3.1.

MiRNAs involved in cancer drug resistance by efflux pathway
hsa-miR-204-3p
hsa-miR-3198
hsa-miR-3618
hsa-miR-432-5p
hsa-miR-4328
hsa-miR-4329
hsa-miR-4435
hsa-miR-4468
hsa-miR-4488
hsa-miR-4499
hsa-miR-4516
hsa-miR-4521
hsa-miR-4522
hsa-miR-4534
hsa-miR-4638-3p
hsa-miR-4654
hsa-miR-4661-3p
hsa-miR-4690-3p
hsa-miR-4693-5p
hsa-miR-4706
hsa-miR-4728-5p
hsa-miR-4740-3p
hsa-miR-4741
hsa-miR-4751
hsa-miR-4753-3p
hsa-miR-4757-5p
hsa-miR-4778-5p
hsa-miR-4781-3p
hsa-miR-4787-5p
hsa-miR-4792
hsa-miR-4793-5p
hsa-miR-4794
hsa-miR-548ab
hsa-miR-548ac
hsa-miR-548aj-3p
hsa-miR-5693
hsa-miR-5694
hsa-miR-9-5p

A.

Top Networks

ID	Associated Network Functions	Score
1	Cell Death and Survival, Cancer, Cell Cycle	39
2	Cell Signaling, Cardiovascular Disease, Dermatological Diseases and Conditions	12
3	Cell Death and Survival, Cellular Development, Cellular Growth and Proliferation	4

B.

Top Tox Functions

Molecular and Cellular Functions

Name	p-value	#Molecules
Cellular Movement	6.54E-05 - 9.58E-18	21
Cell Death and Survival	6.95E-05 - 1.41E-17	21
Cellular Development	6.23E-05 - 2.26E-16	22
Cellular Growth and Proliferation	6.95E-05 - 2.26E-16	21
Cell Morphology	6.95E-05 - 6.37E-15	17

C.

Top Canonical Pathways

Name	p-value	Overlap
Cancer Drug Resistance By Drug Efflux	4.59E-69	51.0 % 25/49
PI3K/AKT Signaling	8.68E-41	16.3 % 20/123
Prostate Cancer Signaling	2.54E-40	20.0 % 19/95
Melanoma Signaling	2.67E-39	30.9 % 17/55
PTEN Signaling	2.75E-38	16.0 % 19/119

D. Overlap of canonical pathways

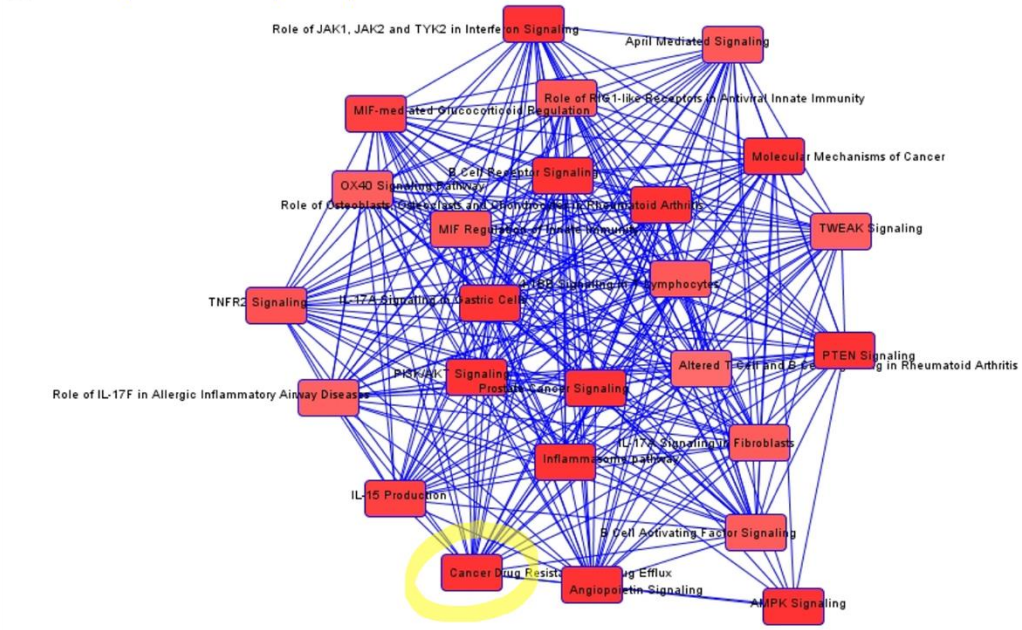
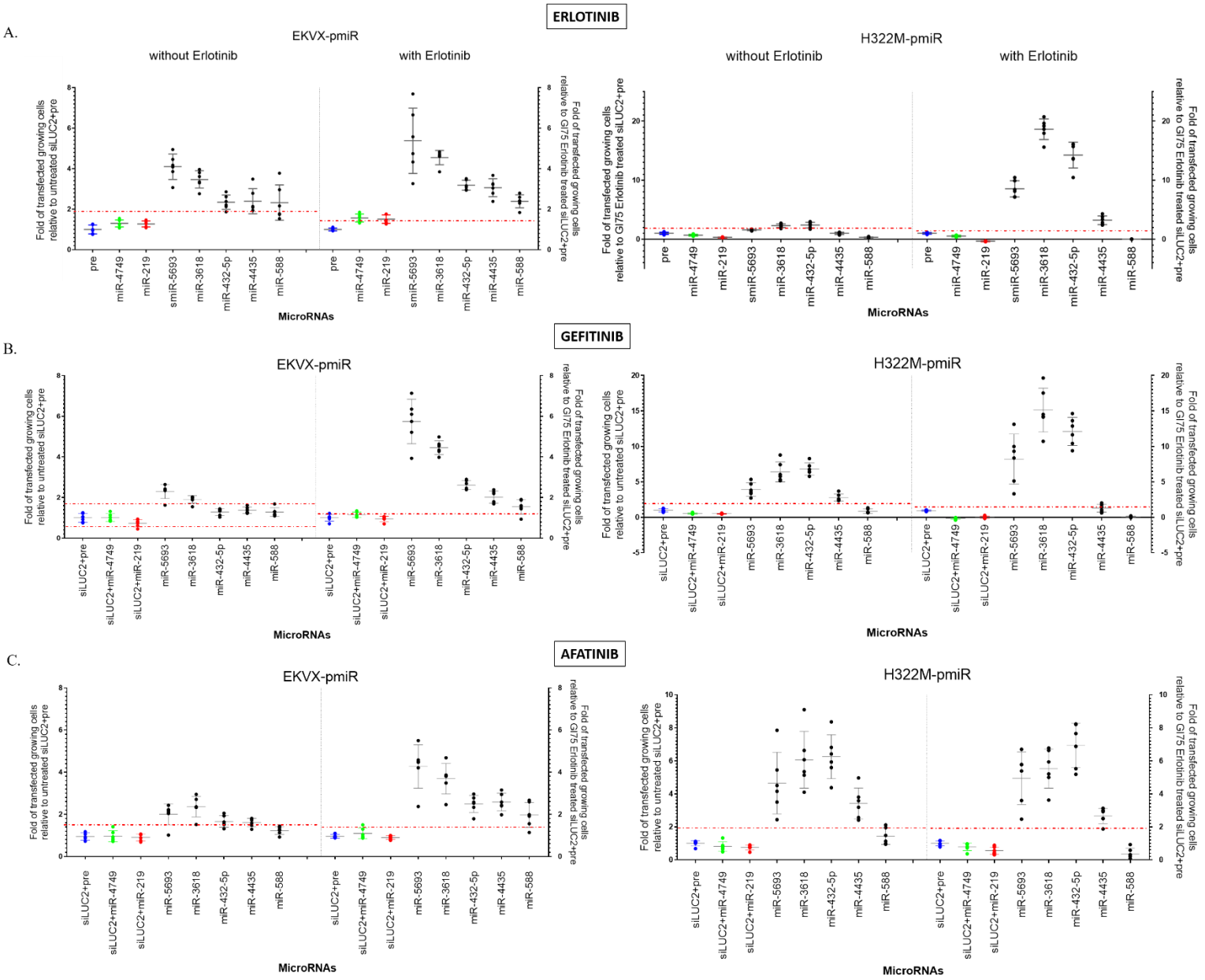


Figure 3.7. Bioinformatic analysis of key pathways regulated by the top 60 miRNAs. Ingenuity Pathway Analysis (IPA) was used to analyze the pathways and cellular functions predicted to be altered by experimentally validated targets of the top 60 miRNAs. A) IPA identified cellular functions altered by the top 60 miRNAs. Score = -log (Fisher's Exact test result calculated by IPA software). Higher the score of a pathway, lower is it probability of being a random event. B) Toxic cellular functions that the targets of the top 60 miRNAs are predicted to significantly dysregulate. C) The top canonical pathways that are significantly altered due to the involvement of the targets of top 60 miRNAs. D) Cross-talk between the canonical pathways identifying novel pathways that may be involved in mediating drug resistance (highlighted in yellow).

3.3.6 Validating erlotinib resistance effects imparted by the top five candidate miRNAs

The top five miRNAs identified as miRNAs involved in sensitizing both EKVX and H322M cells (**Table 3.1)**, were further validated in EKVX-pmiR and H322M-pmiR cells via measurement of luciferase reporter activities, with and without erlotinib treatment. All the newly identified miRNAs enhanced growth of EKVX-pmiR cells in the presence of the erlotinib. Although neither of the positive controls identified in EXVX-pmiR cells promoted resistance in H322M-pmiR cells, four of the newly identified miRNAs promoted resistance (**Figure 3.8A**). Importantly, miRNAs that validated in both cells lines, miR-5693, miR-3618, miR-4435 and miR-432-5p, but not miR-588 were identified by IPA to be regulators of cancer drug resistance and efflux (**Figure 3.8A and Table 3.2**). Apart from erlotinib, it has been reported that certain miRNAs also induce resistance to other TKIs such as Gefitinib (134,139). Therefore, the top 5 miRNAs were evaluated for their ability to enhance growth in the presence of other EGFR-TKIs Afatinib and Gefitinib. Three of five miRNAs, miR-5693, miR-3618 and miR-432-5p consistently enhanced cell growth in the presence of gefitinib in both cell lines (**Figure 3.8B**), but only in EKVX-pmiR cell in the presence of afatinib (**Figure 3.8C**)

Figure 3.8. Validating five miRNAs for promoting resistance to erlotinib and other TKIs. EKVX-pmiR or H322M-pmiR cells were co-transfected with 6nM premiRcontrol (pre) or 6nM of the indicated miRNAs and 0.6nM siLUC2. The positive controls, miR-4749 and miR-219 were also included. Twenty-four hours post-transfection, cells were exposed to GI75 A) erlotinib concentration B) gefitinib C) afatinib or equivalent amount of DMSO (negative control). Seventy-two hours post treatment, reporter activity was measured. Fold growth of transfected cells in the presence or absence of erlotinib are represented relative to negative control (pre). Red dotted lines represent mean plus 4X standard of positive control (miR-4749).


A.
ERLOTINIB
EKVX-pmiR
without Erlotinib
with Erlotinib
H322M-pmiR
without Erlotinib
with Erlotinib
Fold of transfected growing cells relative to untreated siLUC2+pre
Fold of transfected growing cells relative to GI75 Erlotinib treated siLUC2+pre
pre
miR-4749
miR-219
smiR-5693
miR-3618
miR-432-5p
miR-4435
miR-588
MicroRNAs
B.
GEFITINIB
EKVX-pmiR
H322M-pmiR
siLUC2+pre
siLUC2+miR-4749
siLUC2+miR-219
miR-5693
miR-3618
miR-432-5p
miR-4435
miR-588
MicroRNAs
C.
AFATINIB
EKVX-pmiR
H322M-pmiR
MicroRNAs

3.3.7 Level of expression of miR-432 in patient data

For miRNAs with expression data available, expression was evaluated from The Cancer Genome Atlas (TCGA) from both lung adenocarcinoma and lung squamous cell carcinoma (LUAD and LUSC, respectively). MiRNA with high number annotations such as those in the 4000s and 5000s were not present. The heatmap in **Figure 3.9** shows that both miR-204-5p and miR-432-5p are highly expressed in LUAD and LUSC patient samples. While the initial screen identified miR-204 as a miRNA that promoted erlotinib resistance, it was not followed up for two reasons. Firstly, the miRNA that was identified in our screen was miR-204-3p, which is annotated as the passenger strand, while the TCGA data measured the mature strand, or miR-204-5p. Secondarily, the extensive validation studies did not support miR-204-3p in promoting resistance to erlotinib in either of the erlotinib sensitive reporter lines (**Table 3.1**). Hence, based on these findings, we focused on the role of miR-432 as a mediator of erlotinib resistance in NSCLC.

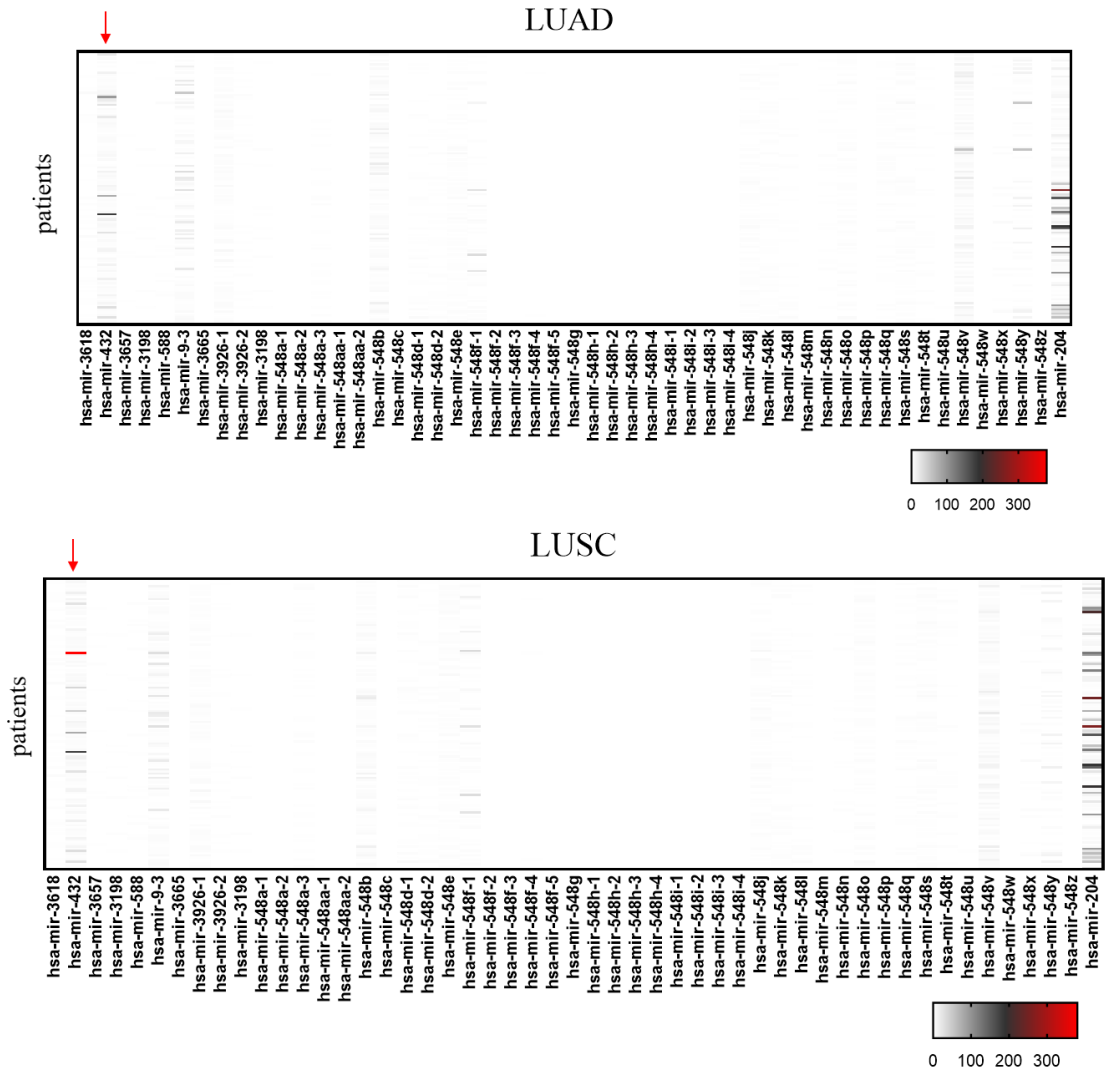
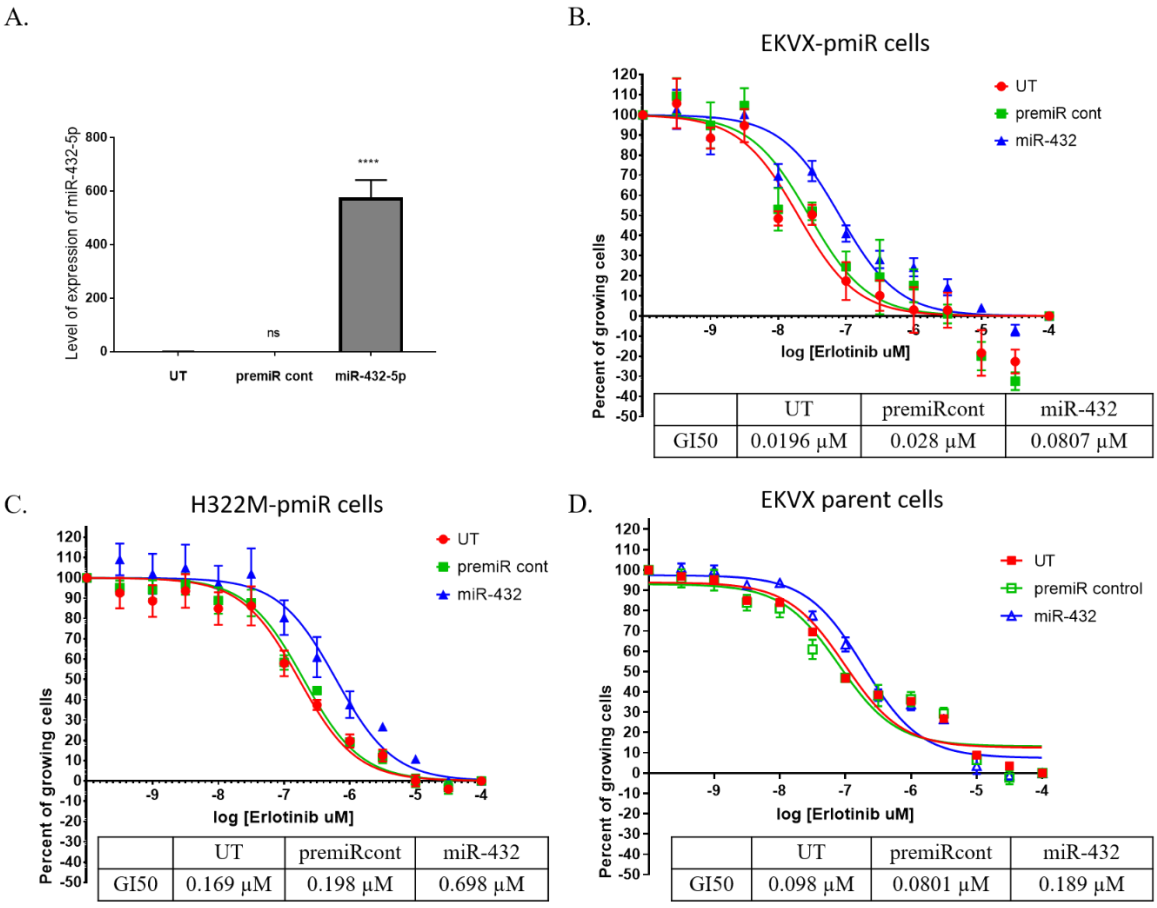


Figure 3.9. Expression of top 60 miRNAs in patient samples. Patient miRNA sequenced data from LUAD and LUSC were retrieved from TCGA. The raw reads per gene value were graphed using GraphPad prism version 7 software to visualize the reads for each miRNA in each patient. The red arrow represents miR-432-5p expression in LUAD and LUSC patients. The scale bars represent the read counts that are undetected or zero (white) baseline (grey) or high (red).

3.3.8 MiR-432-5p mediates development of resistance in erlotinib sensitive NSCLC cells

MiR-432 was transfected into EKVX-pmiR cells, and cell proliferation in the presence of variable concentrations of erlotinib was analyzed via SRB assay (**Figure 3.10 A,B**). This assay was to confirm that luciferase reporter activity of the pmiRGLO plasmid was not off-targeted by miR-432, and that the cells acquired erlotinib resistance upon miR-432 transfection. The growth promoting effect of miR-432 was also evaluated via SRB assay in H322M-pmiR cells, and the

parent EKVX cells (**Figure 3.10 C,D**). In all cases overexpressing miR-432 promoted a significant change in erlotinib sensitivity generating ~2-3 fold increase in GI50 concentrations.



	UT	premiRcont	miR-432
GI50	0.0196 µM	0.028 µM	0.0807 µM

	UT	premiRcont	miR-432
GI50	0.169 µM	0.198 µM	0.698 µM

	UT	premiRcont	miR-432
GI50	0.098 µM	0.0801 µM	0.189 µM

Figure 3.10. MiR-432-5p mediates development of resistance in erlotinib sensitive NSCLC cells. A) Evaluation of miR-432-5p levels in EKVX-pmiR cells post transfection via qRT-PCR. GAPDH was utilized as the endogenous control. One-way ANOVA analysis is used to calculate statistical significance. B) EKVX-pmiR cells , C) H322M-pmiR cells , D) EKVX parental cells were reverse transfected with 6nM premircont or miR-432-5p or were untransfected (UT). Erlotinib dose response via SRB assay was evaluated by exposing cells to varying concentrations of Erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. For percent of cells calculation, number of cells at the time of addition of Erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 erlotinib was calculated from the respective dose curves (as per the NCI-60 Cell Five-Dose Screen, NCI-60, DTP (149).

3.3.9 Antagonizing miR-432-5p does not re-sensitize erlotinib resistant NSCLC cells

In order to evaluate if antagonizing miR-432-5p can rescue the resistant phenotype in NSCLC cells that are endogenously resistant to erlotinib, determined by the NCI-60 DTP study (149), the

following experiments were conducted. Firstly, the endogenous expression of miR-432-5p was evaluated in a panel of NSCLC cells (**Figure 3.11 A**). It was determined that H23 has the highest expression of miR-432-5p relative to parental EKVX cells. Therefore, miR-432-5p was antagonized in H23, levels were confirmed to be downregulated via qRT-PCR (**Figure 3.11 B**), and an SRB assay was conducted in the presence of variable concentrations of erlotinib. Although over 50% of miR-432-5p was reduced via transfection of the antagomir, it did not rescue erlotinib response in H23 cells (**Figure 3.11 C**). Evaluation was conducted on additional erlotinib resistant NSCLC cell lines, H441 and A549, but downregulation of miR-432-5p did not re-sensitize the cells to erlotinib.

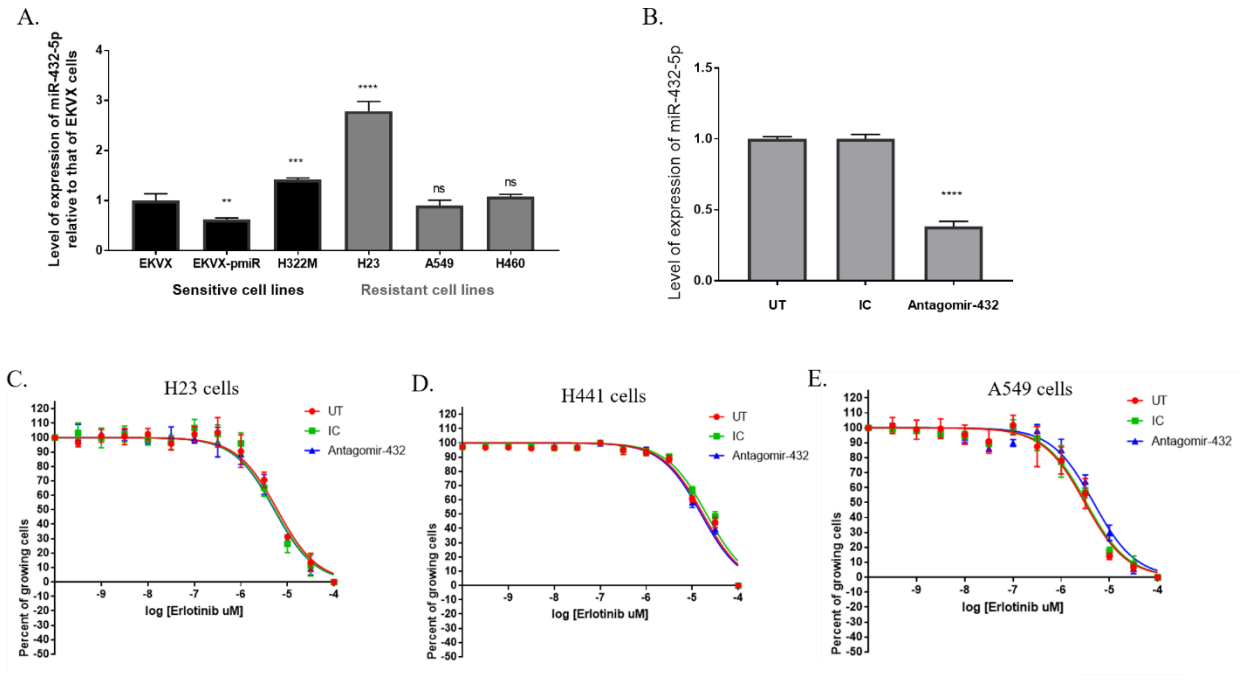


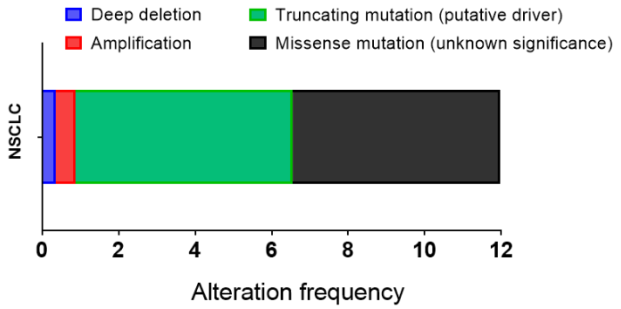
Figure 3.11. Antagonizing miR-432-5p does not re-sensitize erlotinib resistant NSCLC cells. A) Evaluation of endogenous miR-432-5p levels a panel of NSCLC cells, represented relative to EKVX parental cells. B) Evaluation of miR-432-5p levels in H23 cells post antagomiR-432 transfection via qRT-PCR. GAPDH was utilized as the endogenous control for qRT-PCRs. One-way ANOVA analysis is used to calculate statistical significance for A and B. C) H23 cells , D) H441 cells , E) A549 cells were reverse transfected with 150nM inhibitor control (IC) or Antagomir-432-5p or were untransfected (UT). Erlotinib dose response via SRB assay was evaluated by exposing cells to varying concentrations of Erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. For percent of cells calculation, number of cells at the time of addition of Erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 erlotinib was calculated from the respective dose curves (as per the NCI-60 Cell Five-Dose Screen, NCI-60, DTP (149).

3.3.10 Evaluation of putative targets of miR-432-5p that mediate erlotinib resistance

Since antagonizing miR-432-5p is unable to the rescue erlotinib resistant phenotype in NSCLC cells, we hypothesized that miR-432-5p is likely targeting a tumor suppressor gene that is lost in NSCLC cells upon development of resistance. NF1 is one such tumor suppressor gene reported to be severely lost in erlotinib sensitive NSCLC cells, mediating the development of resistance (153). Additionally, cBioPortal (154) analysis of pan-cancer TCGA NSCLC dataset (155) shows that NF1 is altered in 12% of patients, comprising of alterations such as deep deletions (0.35%), amplifications (0.52%), truncating mutations (5.68%), and missense mutations (5.42%). Both deep deletion of NF1, associated with loss of function of NF1 in NSCLC, and truncating mutations in patients are shown to be putative drivers of cancer, and are currently associated to occur with concurrent oncogenic alterations (156,157) (**Figure 3.12 A**). Therefore, bioinformatically, we confirmed that NF1 is a putative target of miR-432-5p via canonical targeting mechanism. Two miR-432-5p targeting sites on the 3'-UTR of *NF1* were predicted by mirmap software (**Figure 3.12 B**), but the site with a higher mirmap score (score = 81.85) was also predicted by TargetScan. TargetScan predicted that the binding site on the 3'-UTR of *NF1* is a poorly conserved but a confidently annotated target of miR-432-5p (**Figure 3.12 C**). Therefore, further evaluation of NF1 post transfection of miR-432-5p in erlotinib sensitive NSCLC cells will ascertain the mechanism of action of miR-432-5p, and confirm if NF1 is a key player in the process.

Additionally, utilizing IPA analysis to identify the pathways regulated by miR-432-5p alone, several metabolic pathways are predicted to be associated with targets of miR-432-5p (**Figure 3.13**). Therefore, it is likely that miR-432-5p mediates erlotinib resistance by triggering metabolic reprogramming in cancer cells (158). Interestingly, a few of the canonical pathways that are significantly associated with targets of miR-432-5p, have been recently reported to be altered in NSCLC resulting in the development of resistance to erlotinib (140) or other TKIs(139).

A.



B.

	miRNA	Gene	ΔG open	Probability exact	Conservation PhyloP	miRmap score
1.	hsa-miR-432-5p	NF1	80.43	30.61	59.03	81.85
2.	hsa-miR-432-5p	NF1	55.43	30.61	73.71	71.88

Predicted site 1

```
                                1056
5' (mRNA)                       |
AUAUUUUUAUUUUGUUAAAUAAUUUCCAAGAAUAGAGUA
                                ||||||
                     GGUGGGUUACUGGAUGAGGUUCU
```

Predicted site 2

```
                                2396
5' (mRNA)                       |
UGACUUUUGAGGAAAAUCUAGCUUUCCAAGUAACUAAAA
                                ||||||
                     GGUGGGUUACUGGAUGAGGUUCU
```

C.

	Predicted consequential pairing of target region (top) and miRNA (bottom)
Position 1050-1056 of NF1 3' UTR	5' ...UUUUGUUAAAUAAUUUCCAAGAA...
	||||||
hsa-miR-432-5p	3' GGUGGGUUACUGGAUGAGGUUCU

Figure 3.12. Bioinformatic analyses of NF1 alteration in NSCLC and predictions of canonical targeting of NF1 via miR-432-5p. A) NF1 gene is frequently altered in NSCLC, accounting for 12% of NSCLC patients in pan-Lung cancer dataset(155), evaluated using cBioPortal (154). Two bioinformatic tools were utilized to identify if miR-432-5p is predicted to target a well-known mediator of erlotinib resistance, *NF1* via the canonical miRNA targeting mechanism (153). B) Two sites on the 3'-UTR of *NF1* were predicted to be targeted by mirmap. C) A single site on *NF1* 3'-UTR was predicted to be targeted by miR-432-5p.

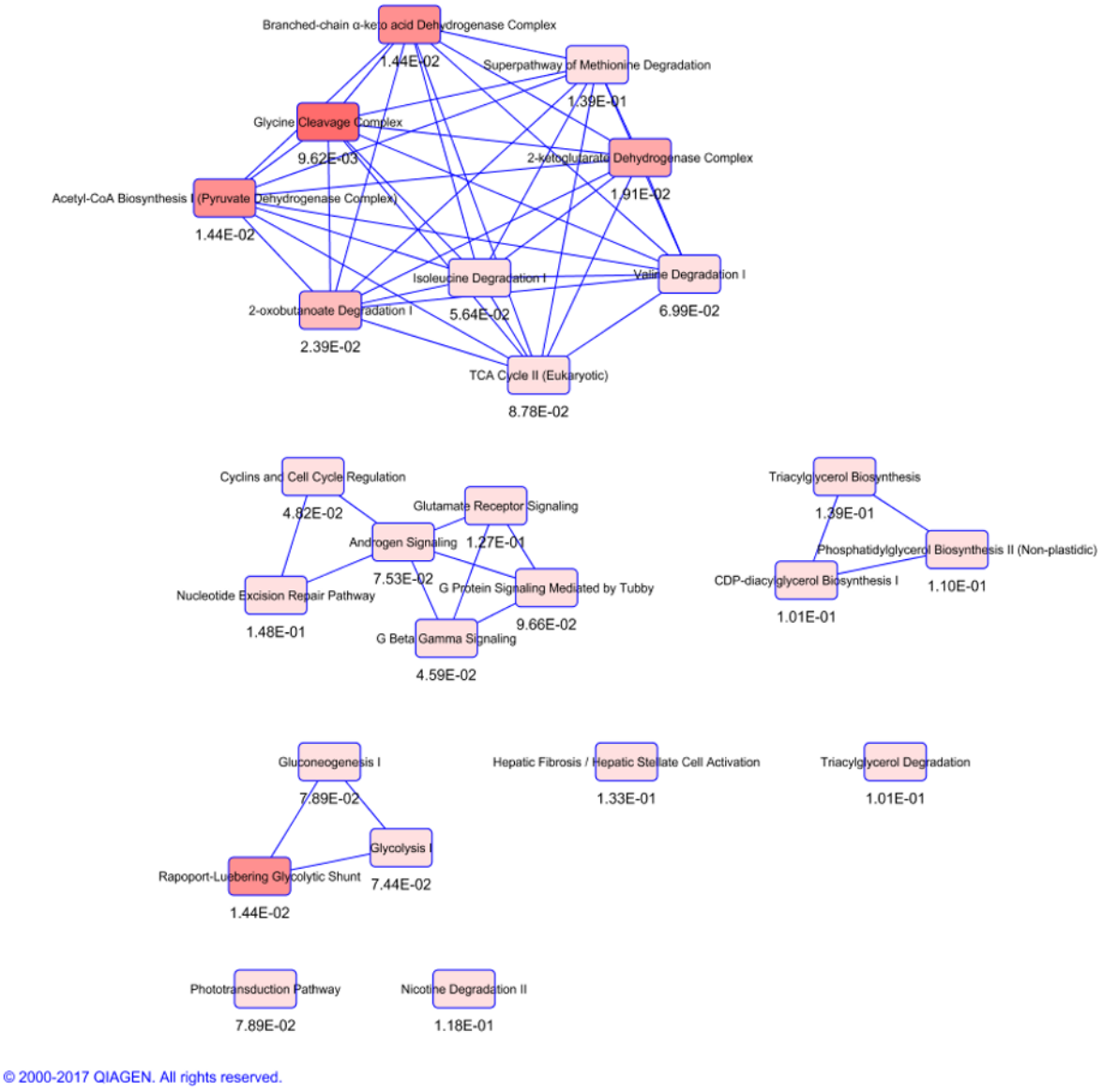


Figure 3.13. Putative pathways regulated by targets of miR-432-5p, predicted by Ingenuity Pathway Analysis (IPA).

3.4 Discussion and future directions:

Currently, there are over 2,656 mature miRNAs annotated in miRbase (version 22)(146). However, functions of over half of the miRNAs remain unknown (146). From the above screen, it can be inferred that several miRNAs are likely to have a role in altering specific phenotypes in cancer, such as resistance to therapies. The results from the overexpression screen, corroborated with the bioinformatic evaluation of targets of the miRNAs being involved in cancer drug

resistance and efflux pathways (results from **Table 3.2**), indicating additional effort is needed to identify how some of these candidate miRNAs mediate resistance.

One of the miRNAs identified in this study, miR-432-5p was found to promote resistance in erlotinib sensitive NSCLC cells (**Figure 3.10**). However, the results support the notion that miR-432-5p is responsible for only the development of resistance (**Figure 3.10**) but antagonizing the miRNA does not result in rescuing the phenotype in resistant cells (**Figure 3.11)**. One potential hypothesis for the failed response to antagonizing miR-432 is that miR-432-5p may be involved in targeting a tumor suppressor gene that is frequently deleted or truncated in lung cancer (156,157), and lost in erlotinib resistant NSCLC cells (153) (**Figure 3.12**). Other possibilities include 1) alternative polyadenylation use of miR-432 targets in resistant NSCLC cells, which is often observed in cancers, resulting in inability of miR-432 targeting, and/or 2) stable changes made to the epigenome that are influenced by miR-432 which cannot be overcome following antagonization.

Therefore, in the future, to validate the findings of this study, the effect of miR-432-5p on the putative target needs to be evaluated. Firstly, experiments need to be conducted to evaluate if miR-432 targets NF1. Depending on the severity of NF1 downregulation, further validation studies will need to be conducted to determine if NF1 is a major facilitator of miR-432-5p mediated erlotinib resistance. If the miR-432-NF1 connection is validated, a direct vs. indirect effect will need to be established, i.e. if the 3'-UTR of NF1 is targeted by miR-432-5p or if targeting is via a non-canonical mechanisms i.e. the 5'-UTR, coding-sequence or the promoter. These mechanisms are explained in more detail in the next chapter. Molecular cloning of the specific regions to evaluate their regulation by miR-432-5p followed by additional cell-based assays such as luciferase assays will confirm the mechanism of NF1 regulation by miR-432-5p.

However, if NF1 does not experimentally identify as a critical facilitator of miR-432-5p mediated erlotinib resistance, evaluation of a significantly associated metabolic pathway predicted by IPA analysis can possibly serve as an alternative approach. Upregulation of branched-chain amino acid (BCAA) pathway leading to reduced reactive oxygen species (ROS) levels, and thereby resulting in erlotinib resistance in NSCLC, is a reported mechanism (140). And since branched chain α-keto acid dehydrogenase complex (BCKDH complex) is a major player in the BCAA pathway,

predicted to be regulated by miR-432-5p (**Figure 3.13**), analysis of enzymes of the BCKDH complex as putative targets of miR-432-5p can be a direction to pursue.

Additionally, generating inducible lines to evaluate if constitutive overexpression of the miR-432-5p generates erlotinib resistant cells, and testing erlotinib response of the cell line with and without induction of miR-432-5p *in vivo* will further validate the findings. Lastly and most importantly, validating if miR-432-5p levels in RNAs isolated from patients post-erlotinib treatment is higher than from matched pre-erlotinib treated patients, will confirm miR-432-5p a mediator of erlotinib resistance.

3.5 References:

1. Crick FH. On protein synthesis. Symp Soc Exp Biol **1958**;12:138-63
2. CRICK F. Central Dogma of Molecular Biology. Nature **1970**;227:561-3
3. Holley RW, Apgar J, Everett GA, Madison JT, Marquisee M, Merrill SH, *et al.* Structure of a Ribonucleic Acid. **1965**
4. Fox GE, Woese CR. The architecture of 5S rRNA and its relation to function. J Mol Evol **1975**;6:61-76
5. Lerner MR, Steitz JA. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A **1979**;76:5495-9
6. Filipowicz W, Kiss T. Structure and function of nucleolar snRNPs. Mol Biol Rep **1993**;18:149-56
7. Cech TR, Steitz JA. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. Cell **2014**;157:77-94
8. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell **1993**;75:843-54
9. Lau NC, Lim LP, Weinstein EG, Bartel DP. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans. **2001**
10. Lee RC, Ambros V. An Extensive Class of Small RNAs in Caenorhabditis elegans. **2001**
11. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of Novel Genes Coding for Small Expressed RNAs. **2001**

12. Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell **1993**;75:855-62
13. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature **2000**;403:901-6
14. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, *et al.* Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature **2000**;408:86-9
15. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res **2014**;42:D68-73
16. Hausser J, Zavolan M. Identification and consequences of miRNA-target interactions [mdash] beyond repression of gene expression. Nature Reviews Genetics **2014**;15:599-612
17. Cipolla GA. A non-canonical landscape of the microRNA system. Front Genet **2014**;5:337
18. Orellana EA, Kasinski AL. MicroRNAs in Cancer: A Historical Perspective on the Path from Discovery to Therapy. Cancers (Basel) **2015**;7:1388-405
19. Steinkraus BR, Radcliffe Department of Medicine UoOWIoMMOU, Toegel M, Radcliffe Department of Medicine UoOWIoMMOU, Fulga TA, Radcliffe Department of Medicine UoOWIoMMOU. Tiny giants of gene regulation: experimental strategies for microRNA functional studies. Wiley Interdisciplinary Reviews: Developmental Biology **2016**;5:311-62
20. Rupaimoole R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nature Reviews Drug Discovery **2017**;16:203-22
21. Kasinski AL, Slack FJ. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. Nature Reviews Cancer **2011**;11:849-64
22. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. Nat Rev Cancer **2015**;15:321-33
23. Calin GA, Fabbri M. 4 – Epigenetics and miRNAs in Human Cancer. **2010**;70:87–99
24. Sato F, Tsuchiya S, Meltzer SJ, Shimizu K. MicroRNAs and epigenetics. Febs j **2011**;278:1598-609

25. Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N, *et al.* Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. Mol Cell **2007**;26:731-43
26. Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, *et al.* A global map of p53 transcription-factor binding sites in the human genome. Cell **2006**;124:207-19
27. Hermeking H. p53 enters the microRNA world. Cancer Cell **2007**;12:414-8
28. Fabbri M, Bottoni A, Shimizu M, Spizzo R, Nicoloso MS, Rossi S, *et al.* Association of a microRNA/TP53 feedback circuitry with pathogenesis and outcome of B-cell chronic lymphocytic leukemia. Jama **2011**;305:59-67
29. Kim T, Veronese A, Pichiorri F, Lee TJ, Jeon YJ, Volinia S, *et al.* p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. J Exp Med **2011**;208:875-83
30. Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, Li CW, *et al.* p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. Nat Cell Biol **2011**;13:317-23
31. Wang Z, Lin S, Li JJ, Xu Z, Yao H, Zhu X, *et al.* MYC protein inhibits transcription of the microRNA cluster MC-let-7a-1~let-7d via noncanonical E-box. J Biol Chem **2011**;286:39703-14
32. Chang TC, Yu D, Lee YS, Wentzel EA, Arking DE, West KM, *et al.* Widespread microRNA repression by Myc contributes to tumorigenesis. Nat Genet **2008**;40:43-50
33. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature **2005**;435:839-43
34. Schulte JH, Horn S, Otto T, Samans B, Heukamp LC, Eilers UC, *et al.* MYCN regulates oncogenic MicroRNAs in neuroblastoma. Int J Cancer **2008**;122:699-704
35. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, *et al.* Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res **2007**;67:1424-9
36. Saito Y, Liang G, Egger G, Friedman JM, Chuang JC, Coetzee GA, *et al.* Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell **2006**;9:435-43

37. Nadal E, Chen G, Gallegos M, Lin L, Ferrer-Torres D, Truini A, *et al.* Epigenetic inactivation of microRNA-34b/c predicts poor disease-free survival in early-stage lung adenocarcinoma. Clin Cancer Res **2013**;19:6842-52
38. Vogt M, Munding J, Gruner M, Liffers ST, Verdoodt B, Hauk J, *et al.* Frequent concomitant inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal, pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft tissue sarcomas. Virchows Arch **2011**;458:313-22
39. Wang Z, Chen Z, Gao Y, Li N, Li B, Tan F, *et al.* DNA hypermethylation of microRNA-34b/c has prognostic value for stage non-small cell lung cancer. Cancer Biol Ther **2011**;11:490-6
40. Saito Y, Friedman JM, Chihara Y, Egger G, Chuang JC, Liang G. Epigenetic therapy upregulates the tumor suppressor microRNA-126 and its host gene EGFL7 in human cancer cells. Biochem Biophys Res Commun **2009**;379:726-31
41. Bohlig L, Friedrich M, Engeland K. p53 activates the PANK1/miRNA-107 gene leading to downregulation of CDK6 and p130 cell cycle proteins. Nucleic Acids Res **2011**;39:440-53
42. Zhang X, Zeng Y. Regulation of mammalian microRNA expression. J Cardiovasc Transl Res **2010**;3:197-203
43. Brueckner B, Stresemann C, Kuner R, Mund C, Musch T, Meister M, *et al.* The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. Cancer Res **2007**;67:1419-23
44. Bandres E, Agirre X, Bitarte N, Ramirez N, Zarate R, Roman-Gomez J, *et al.* Epigenetic regulation of microRNA expression in colorectal cancer. Int J Cancer **2009**;125:2737-43
45. Li X, Carthew RW. A microRNA Mediates EGF Receptor Signaling and Promotes Photoreceptor Differentiation in the Drosophila Eye. Cell **2005**;123:1267-77
46. Hirata Y, Murai N, Yanaihara N, Saito M, Urashima M, Murakami Y, *et al.* MicroRNA-21 is a candidate driver gene for 17q23-25 amplification in ovarian clear cell carcinoma. BMC Cancer **2014**;14:799
47. Buscaglia LE, Li Y. Apoptosis and the target genes of microRNA-21. Chin J Cancer **2011**;30:371-80

48. Poliseno L, Salmena L, Riccardi L, Fornari A, Song MS, Hobbs RM, *et al.* Identification of the miR-106b~25 MicroRNA Cluster as a Proto-Oncogenic PTEN-Targeting Intron That Cooperates with Its Host Gene MCM7 in Transformation. **2010**
49. Mehlich D, Garbicz F, Wlodarski PK. The emerging roles of the polycistronic miR-106b approximately 25 cluster in cancer - A comprehensive review. Biomed Pharmacother **2018**;107:1183-95
50. Landais S, Landry S, Legault P, Rassart E. Oncogenic Potential of the miR-106-363 Cluster and Its Implication in Human T-Cell Leukemia. **2007**
51. Li M, Zhou Y, Xia T, Zhou X, Huang Z, Zhang H, *et al.* Circulating microRNAs from the miR-106a–363 cluster on chromosome X as novel diagnostic biomarkers for breast cancer. Breast Cancer Res Treat2018. p 257-70.
52. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, *et al.* Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A **2004**;101:2999-3004
53. Lerner M, Harada M, Loven J, Castro J, Davis Z, Oscier D, *et al.* DLEU2, frequently deleted in malignancy, functions as a critical host gene of the cell cycle inhibitory microRNAs miR-15a and miR-16-1. Exp Cell Res **2009**;315:2941-52
54. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, *et al.* MicroRNA genes are transcribed by RNA polymerase II. Embo j **2004**;23:4051-60
55. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell **2004**;116:281-97
56. Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, *et al.* Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell **2006**;125:887-901
57. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, *et al.* The Microprocessor complex mediates the genesis of microRNAs. Nature **2004**;432:235-40
58. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev **2003**;17:3011-6
59. Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, *et al.* Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature **2011**;475:201-5

60. Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature **2005**;436:740-4
61. Kawamata T, Yoda M, Tomari Y. Multilayer checkpoints for microRNA authenticity during RISC assembly. EMBO Rep2011. p 944-9.
62. Zhang X, Niu D, Carbonell A, Wang A, Lee A, Tun V, *et al.* ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. Nature Communications, Published online: 19 November 2014; | doi:101038/ncomms6468 **2014**
63. Cenik ES, Zamore PD. Argonaute proteins. Curr Biol **2011**;21:R446-9
64. Burroughs AM, Ando Y, de Hoon MJ, Tomaru Y, Suzuki H, Hayashizaki Y, *et al.* Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. RNA Biol **2011**;8:158-77
65. Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. Science **2004**;305:1437-41
66. Dueck A, Ziegler C, Eichner A, Berezikov E, Meister G. microRNAs associated with the different human Argonaute proteins. Nucleic Acids Res **2012**;40:9850-62
67. Nakanishi K. Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? Wiley Interdiscip Rev RNA2016. p 637-60.
68. Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, *et al.* New class of microRNA targets containing simultaneous 5′-UTR and 3′-UTR interaction sites. Genome Res **2009**;19:1175-83
69. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell **2009**;136:215-33
70. Doench JG, Petersen CP, Sharp PA. siRNAs can function as miRNAs. Genes Dev **2003**;17:438-42
71. Zhang JG, Wang JJ, Zhao F, Liu Q, Jiang K, Yang GH. MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). Clin Chim Acta **2010**;411:846-52

72. Lu Z, Liu M, Stribinskis V, Klinge CM, Ramos KS, Colburn NH, *et al.* MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. Oncogene **2008**;27:4373-9
73. Ha M, Kim VN. Regulation of microRNA biogenesis. Nature Reviews Molecular Cell Biology **2014**;15:509-24
74. Zhao J, Tao Y, Zhou Y, Qin N, Chen C, Tian D, *et al.* MicroRNA-7: a promising new target in cancer therapy. Cancer Cell Int **2015**;15:103
75. Kefas B, Godlewski J, Comeau L, Li Y, Abounader R, Hawkinson M, *et al.* microRNA-7 Inhibits the Epidermal Growth Factor Receptor and the Akt Pathway and Is Down-regulated in Glioblastoma. **2008**
76. Roush S, Slack FJ. The let-7 family of microRNAs. Trends Cell Biol **2008**;18:505-16
77. Lee H, Han S, Kwon CS, Lee D. Biogenesis and regulation of the let-7 miRNAs and their functional implications. Protein Cell **2016**;7:100-13
78. Balca-Silva J, Sousa Neves S, Goncalves AC, Abrantes AM, Casalta-Lopes J, Botelho MF, *et al.* Effect of miR-34b overexpression on the radiosensitivity of non-small cell lung cancer cell lines. Anticancer Res **2012**;32:1603-9
79. Tanaka N, Toyooka S, Soh J, Kubo T, Yamamoto H, Maki Y, *et al.* Frequent methylation and oncogenic role of microRNA-34b/c in small-cell lung cancer. Lung Cancer **2012**;76:32-8
80. Lize M, Herr C, Klimke A, Bals R, Dobbelstein M. MicroRNA-449a levels increase by several orders of magnitude during mucociliary differentiation of airway epithelia. Cell Cycle **2010**;9:4579-83
81. Bou Kheir T, Futoma-Kazmierczak E, Jacobsen A, Krogh A, Bardram L, Hother C, *et al.* miR-449 inhibits cell proliferation and is down-regulated in gastric cancer. Mol Cancer **2011**;10:29
82. Mao A, Zhao Q, Zhou X, Sun C, Si J, Zhou R, *et al.* MicroRNA-449a enhances radiosensitivity by downregulation of c-Myc in prostate cancer cells. Sci Rep **2016**;6:27346
83. Concepcion CP, Han YC, Mu P, Bonetti C, Yao E, D'Andrea A, *et al.* Intact p53-dependent responses in miR-34-deficient mice. PLoS Genet **2012**;8:e1002797

84. Comazzetto S, Di Giacomo M, Rasmussen KD, Much C, Azzi C, Perlas E, *et al.* Oligoasthenoteratozoospermia and infertility in mice deficient for miR-34b/c and miR-449 loci. PLoS Genet **2014**;10:e1004597
85. Song R, Walentek P, Sponer N, Klimke A, Lee JS, Dixon G, *et al.* miR-34/449 miRNAs are required for motile ciliogenesis by repressing cp110. Nature **2014**;510:115-20
86. Sun J, Gao B, Zhou M, Wang ZZ, Zhang F, Deng JE, *et al.* Comparative genomic analysis reveals evolutionary characteristics and patterns of microRNA clusters in vertebrates. Gene **2013**;512:383-91
87. Khuu C, Utheim TP, Sehic A. The Three Paralogous MicroRNA Clusters in Development and Disease, miR-17-92, miR-106a-363, and miR-106b-25. Scientifica (Cairo) **2016**;2016
88. Hertel J, Lindemeyer M, Missal K, Fried C, Tanzer A, Flamm C, *et al.* The expansion of the metazoan microRNA repertoire. BMC Genomics **2006**;7:25
89. Mogilyansky E, Rigoutsos I. The miR-17|[sol]|92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell Death & Differentiation **2013**;20:1603-14
90. Ventura A, Young AG, Winslow MM, Lintault L, Meissner A, Erkeland SJ, *et al.* Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. Cell **2008**;132:875-86
91. Mu P, Han YC, Betel D, Yao E, Squatrito M, Ogrodowski P, *et al.* Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. Genes Dev **2009**;23:2806-11
92. Esquela-Kerscher A, Slack FJ. Oncomirs |[mdash]| microRNAs with a role in cancer. Nature Reviews Cancer **2006**;6:259-69
93. Xiao C, Srinivasan L, Calado DP, Patterson HC, Zhang B, Wang J, *et al.* Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. Nat Immunol **2008**;9:405-14
94. Olive V, Jiang I, He L. mir-17-92, a cluster of miRNAs in the midst of the cancer network. Int J Biochem Cell Biol **2010**;42:1348-54

95. Olive V, Sabio E, Bennett MJ, De Jong CS, Biton A, McGann JC, *et al.* A component of the mir-17-92 polycistronic oncomir promotes oncogene-dependent apoptosis. Elife **2013**;2:e00822
96. Olive V, Bennett MJ, Walker JC, Ma C, Jiang I, Cordon-Cardo C, *et al.* miR-19 is a key oncogenic component of mir-17-92. Genes Dev **2009**;23:2839-49
97. Sandhu SK, Fassan M, Volinia S, Lovat F, Balatti V, Pekarsky Y, *et al.* B-cell malignancies in microRNA Eμ-miR-17∼92 transgenic mice. Proc Natl Acad Sci U S A2013. p 18208-13.
98. Du P, Wang L, Sliz P, Gregory RI. A Biogenesis Step Upstream of Microprocessor Controls miR-17 approximately 92 Expression. Cell **2015**;162:885-99
99. Donayo AO, Johnson RM, Tseng HW, Izreig S, Gariepy A, Mayya VK, *et al.* Oncogenic Biogenesis of pri-miR-17 approximately 92 Reveals Hierarchy and Competition among Polycistronic MicroRNAs. Mol Cell **2019**;75:340-56.e10
100. Medina PP, Nolde M, Slack FJ. OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature **2010**;467:86-U119
101. Babar IA, Cheng CJ, Booth CJ, Liang X, Weidhaas JB, Saltzman WM, *et al.* Nanoparticle-based therapy in an in vivo microRNA-155 (miR-155)-dependent mouse model of lymphoma. Proceedings of the National Academy of Sciences of the United States of America **2012**;109:E1695-E704
102. Costinean S, Zanesi N, Pekarsky Y, Tili E, Volinia S, Heerema N, *et al.* Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proc Natl Acad Sci U S A **2006**;103:7024-9
103. Costinean S, Sandhu SK, Pedersen IM, Tili E, Trotta R, Perrotti D, *et al.* Src homology 2 domain-containing inositol-5-phosphatase and CCAAT enhancer-binding protein beta are targeted by miR-155 in B cells of Emicro-MiR-155 transgenic mice. Blood **2009**;114:1374-82
104. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, *et al.* RAS is regulated by the let-7 microRNA family. Cell **2005**;120:635-47
105. Kumar MS, Erkeland SJ, Pester RE, Chen CY, Ebert MS, Sharp PA, *et al.* Suppression of non-small cell lung tumor development by the let-7 microRNA family. Proc Natl Acad Sci U S A **2008**;105:3903-8

106. Yu F, Yao H, Zhu P, Zhang X, Pan Q, Gong C, *et al.* let-7 regulates self renewal and tumorigenicity of breast cancer cells. Cell **2007**;131:1109-23
107. Newman MA, Thomson JM, Hammond SM. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. Rna **2008**;14:1539-49
108. Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell **2008**;32:276-84
109. Viswanathan SR, Daley GQ, Gregory RI. Selective blockade of microRNA processing by Lin28. Science **2008**;320:97-100
110. Piskounova E, Viswanathan SR, Janas M, LaPierre RJ, Daley GQ, Sliz P, *et al.* Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J Biol Chem **2008**;283:21310-4
111. Sampson VB, Rong NH, Han J, Yang Q, Aris V, Soteropoulos P, *et al.* MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. Cancer Res **2007**;67:9762-70
112. Misso G, Di Martino MT, De Rosa G, Farooqi AA, Lombardi A, Campani V, *et al.* Mir-34: a new weapon against cancer? Mol Ther Nucleic Acids **2014**;3:e194
113. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, *et al.* Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A2002. p 15524-9.
114. Tsujimoto Y, Finger LR, Yunis J, Nowell PC, Croce CM. Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. Science **1984**;226:1097-9
115. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A **2005**;102:13944-9
116. Bonci D, Coppola V, Musumeci M, Addario A, Giuffrida R, Memeo L, *et al.* The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. Nat Med **2008**;14:1271-7
117. Calin GA, Cimmino A, Fabbri M, Ferracin M, Wojcik SE, Shimizu M, *et al.* MiR-15a and miR-16-1 cluster functions in human leukemia. Proc Natl Acad Sci U S A **2008**;105:5166-71

118. Pekarsky Y, Croce CM. Role of miR-15/16 in CLL. Cell Death Differ **2015**;22:6-11
119. Aqeilan RI, Calin GA, Croce CM. miR-15a and miR-16-1 in cancer: discovery, function and future perspectives. Cell Death Differ **2010**;17:215-20
120. Luo SY, Lam DC. Oncogenic driver mutations in lung cancer. Translational Respiratory Medicine **2013**;1:1-8
121. Seo JS, Ju YS, Lee WC, Shin JY, Lee JK, Bleazard T, *et al.* The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res **2012**;22:2109-19
122. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, *et al.* MET Amplification Leads to Gefitinib Resistance in Lung Cancer by Activating ERBB3 Signaling. **2007**
123. Soria JC, Lee HY, Lee JI, Wang L, Issa JP, Kemp BL, *et al.* Lack of PTEN expression in non-small cell lung cancer could be related to promoter methylation. Clin Cancer Res **2002**;8:1178-84
124. Gusenbauer S, Vlaicu P, Ullrich A. HGF induces novel EGFR functions involved in resistance formation to tyrosine kinase inhibitors. Oncogene **2012**;32:3846-56
125. Frampton GM, Ali SM, Rosenzweig M, Chmielecki J, Lu X, Bauer TM, *et al.* Activation of MET via Diverse Exon 14 Splicing Alterations Occurs in Multiple Tumor Types and Confers Clinical Sensitivity to MET Inhibitors. **2015**
126. Forgacs E, Biesterveld EJ, Sekido Y, Fong K, Muneer S, Wistuba II, *et al.* Mutation analysis of the PTEN/MMAC1 gene in lung cancer. Oncogene **1998**;17:1557-65
127. Kasinski AL, Kelnar K, Stahlhut C, Orellana E, Zhao J, Shimer E, *et al.* A combinatorial microRNA therapeutics approach to suppressing non-small cell lung cancer. Oncogene **2015**;34:3547-55
128. Bandi N, Vassella E. miR-34a and miR-15a/16 are co-regulated in non-small cell lung cancer and control cell cycle progression in a synergistic and Rb-dependent manner. Mol Cancer **2011**;10:55
129. Li Y-L, Liu X-M, Zhang C-Y, Zhou J-B, Shao Y, Liang C, *et al.* MicroRNA-34a/EGFR axis plays pivotal roles in lung tumorigenesis. Oncogenesis **2017**;6
130. Yang Y, Meng H, Peng Q, Yang X, Gan R, Zhao L, *et al.* Downregulation of microRNA-21 expression restrains non-small cell lung cancer cell proliferation and migration through upregulation of programmed cell death 4. Cancer Gene Ther **2015**;22:23-9

131. Donnem T, Eklo K, Berg T, Sorbye SW, Lonvik K, Al-Saad S, *et al.* Prognostic Impact of MiR-155 in Non-Small Cell Lung Cancer Evaluated by in Situ Hybridization. J Transl Med2011. p 6.
132. Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, *et al.* A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res **2005**;65:9628-32
133. Yang M, Shen H, Qiu C, Ni Y, Wang L, Dong W, *et al.* High expression of miR-21 and miR-155 predicts recurrence and unfavourable survival in non-small cell lung cancer. Eur J Cancer **2013**;49:604-15
134. Shen H, Zhu F, Liu J, Xu T, Pei D, Wang R, *et al.* Alteration in Mir-21/PTEN expression modulates gefitinib resistance in non-small cell lung cancer. PLoS One **2014**;9:e103305
135. Bing Li SR, Xuefei Li, Yongsheng Wanga, David Garfield, Songwen Zhoua, Xiaoxia Chena, Chunxia Sua, Mo Chena, Peng Kuanga, Guanghui Gaoa, Yayi Hea, Lihong Fana, Ke Feia, Caicun Zhoua, Gerald Schmit-Bindertd. MiR-21 overexpression is associated with acquired resistance of EGFR-TKI in non-small cell lung cancer. **2014**;83:146–53
136. Zhou JY, Chen X, Zhao J, Bao Z, Zhang P, Liu ZF. MicroRNA-34a overcomes HGF-mediated gefitinib resistance in EGFR mutant lung cancer cells partly by targeting MET. Cancer Lett **2014**;351:265-71
137. Stahlhut C, Slack FJ. Combinatorial Action of MicroRNAs let-7 and miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung Cancer Cell Proliferation. https://doiorg/101080/1538410120141003008 **2015**
138. Zhao J, Kelnar K, Bader AG. In-depth analysis shows synergy between erlotinib and miR-34a. PLoS One **2014**;9:e89105
139. Zhang WC, Wells JM, Chow KH, Huang H, Yuan M, Saxena T, *et al.* miR-147b-mediated TCA cycle dysfunction and pseudohypoxia initiate drug tolerance to EGFR inhibitors in lung adenocarcinoma. Nat Metab **2019**;1:460-74
140. Wang Y, Zhang J, Ren S, Sun D, Huang HY, Wang H, *et al.* Branched-Chain Amino Acid Metabolic Reprogramming Orchestrates Drug Resistance to EGFR Tyrosine Kinase Inhibitors. Cell Rep **2019**;28:512-25.e6

141. Momcilovic M, Bailey ST, Lee JT, Fishbein MC, Magyar C, Braas D, *et al.* Targeted Inhibition of EGFR and Glutaminase Induces Metabolic Crisis in EGFR Mutant Lung Cancer. Cell Rep **2017**;18:601-10
142. Zhao X, Lu C, Chu W, Zhang B, Zhen Q, Wang R, *et al.* MicroRNA-124 suppresses proliferation and glycolysis in non–small cell lung cancer cells by targeting AKT–GLUT1/HKII:. http://dxdoiorg/101177/1010428317706215 **2017**
143. Lin C, Xie L, Lu Y, Hu Z, Chang J. miR-133b reverses cisplatin resistance by targeting GSTP1 in cisplatin-resistant lung cancer cells. Int J Mol Med **2018**;41:2050-8
144. Zhang X, Zhu J, Xing R, Tie Y, Fu H, Zheng X, *et al.* miR-513a-3p sensitizes human lung adenocarcinoma cells to chemotherapy by targeting GSTP1. Lung Cancer **2012**;77:488-94
145. Pal AS, Kasinski AL. Animal Models to Study MicroRNA Function. Adv Cancer Res **2017**;135:53-118
146. Kozomara A, School of Biological Sciences FoB, Medicine and Health, University of Manchester, Manchester M13 9PT, UK, Birgaoanu M, School of Biological Sciences FoB, Medicine and Health, University of Manchester, Manchester M13 9PT, UK, Griffiths-Jones S, School of Biological Sciences FoB, Medicine and Health, University of Manchester, Manchester M13 9PT, UK. miRBase: from microRNA sequences to function. Nucleic Acids Research **2020**;47
147. 2015 TargetScanHuman 5.2 Custom. <http://www.targetscan.org/vert_50/seedmatch.html>.
148. mirmap. 2020 miRmap web. <https://mirmap.ezlab.org/app/>.
149. NCI-60 D. 2020 NCI-60 Human Tumor Cell Lines Screen | Discovery & Development Services | Developmental Therapeutics Program (DTP). <https://dtp.cancer.gov/discovery_development/nci-60/>.
150. Zhang W, Lin J, Wang P, Sun J. miR-17-5p down-regulation contributes to erlotinib resistance in non-small cell lung cancer cells. J Drug Target **2017**;25:125-31
151. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, *et al.* Genotypic and Histological Evolution of Lung Cancers Acquiring Resistance to EGFR Inhibitors. **2011**

152. Niederst MJ, Sequist LV, Poirier JT, Mermel CH, Lockerman EL, Garcia AR, *et al.* RB loss in resistant EGFR mutant lung adenocarcinomas that transform to small-cell lung cancer. Nat Commun **2015**;6:6377
153. de Bruin EC, Cowell C, Warne PH, Jiang M, Saunders RE, Melnick MA, *et al.* Reduced NF1 expression confers resistance to EGFR inhibition in lung cancer. Cancer Discov **2014**;4:606-19
154. 2020 cBioPortal for Cancer Genomics: NF1 in NSCLC (TCGA 2016). <https://www.cbioportal.org/ >.
155. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. Nature Genetics **2013**;45:1113-20
156. Redig AJ, Capelletti M, Dahlberg SE, Sholl LM, Mach S, Fontes C, *et al.* Clinical and Molecular Characteristics of NF1-Mutant Lung Cancer. **2016**
157. Hayashi T, Desmeules P, Smith RS, Drilon A, Somwar R, Ladanyi M. RASA1 and NF1 are Preferentially Co-Mutated and Define A Distinct Genetic Subset of Smoking-Associated Non-Small Cell Lung Carcinomas Sensitive to MEK Inhibition. Clin Cancer Res **2018**;24:1436-47
158. Beloueche-Babari M, Box C, Arunan V, Parkes HG, Valenti M, De Haven Brandon A, *et al.* Acquired resistance to EGFR tyrosine kinase inhibitors alters the metabolism of human head and neck squamous carcinoma cells and xenograft tumours. Br J Cancer2015. p 1206-14.

CHAPTER 4. IDENTIFICATION OF NOVEL MICRORNA-TARGET PARTNERS AS MEDIATORS OF ERLOTINIB RESISTANCE IN NON SMALL CELL LUNG CANCER

A part of the introduction of following chapter has been adapted from a first author publication. Pal, Arpita S, and Andrea L Kasinski. "Animal Models to Study MicroRNA Function." Advances in cancer research vol. 135 (2017): 53-118. doi:10.1016/bs.acr.2017.06.006
(License (License # 4744250655345) for reuse in thesis/dissertation, for both print and electronic formats are provided by Elsevier and Copyright Clearance Center)

Chapter Overview

In this chapter, the processes involved in the identification of miRNAs and their putative targets are explained. Two microRNAs were identified that promote erlotinib resistance from the overexpression screen (described in Chapter 3), and their putative targets were identified from the knock-out screen (described in Chapter 2). The predicted interactions were functionally validated, and future directions for this study are laid out here.

4.1 Introduction:

Genes that code for proteins are regulated by multiple mechanisms in a cell, and several RNAs, known as non-coding RNAs play critical roles in this regulation (1–6). One such mechanism of regulation is exerted by microRNAs (miRNAs) which are small non-coding RNA molecules ranging between 20-23 nucleotides in length. MiRNAs function by inhibiting translation of their targets either though translational repression or target degradation. Regardless of the mechanism of target gene down regulation, miRNAs engage in the regulation of multiple normal cellular processes such as growth, proliferation, survival, and cell death (7,8). However, in cancers, miRNAs have been reported to become globally dysregulated, thereby manifesting various hallmarks of cancers, such as resistance to cancer therapies by dysregulating proteins encoded by the target transcripts (7,8). In order to study the specific role of a miRNA or a protein in the development of a certain cancer phenotype, their levels are altered, and the resultant phenotype is evaluated. In this study, miRNAs and proteins were altered individually to evaluate their role as critical mediators of resistance to erlotinib, a targeted therapy used to treat EGFR-dependent non-small cell lung cancer (NSCLC) patients. Using both a protein-coding gene knockout approach

(Chapter 2) and a miRNA overexpression strategy (Chapter 3), an intersection between the two approaches was identified.

Currently in the field, miRNAs can be conditionally upregulated, downregulated or knocked-out by many ways, both in cells and *in vivo* (**Figure 4.1, 4.2**) (8). In the study conducted and described in Chapter 3, the most common and transient method to overexpress a miRNA was employed to elucidate the contribution that individual miRNAs have to the development of erlotinib resistance. In this case, over 2,000 human encoded miRNA mimics were overexpressed in erlotinib sensitive cells to identify miRNAs that promote resistance. MiRNA mimics when transfected into cells, act as mature miRNAs. Certain miRNAs that function as oncogenic miRNAs in a cancer cell, when exogenously introduced at high levels as miRNA mimics, can potentially function as an oncomiR resulting in oncogenic transformation of the cell (7,9,10). On the other hand, to study the specific role of a protein coding gene in the development of cancer phenotypes, that gene can essentially be knocked-out by any of the mechanisms described for miRNA genes in **Figure 4.1 (II)** or conditionally overexpressed as illustrated in **Figure 4.2**. In the study presented in Chapter 2, to evaluate loss of function of a protein in mediating development of erlotinib resistance, the CRISPR-Cas9 system was utilized as described in **Figure 4.1 (II) D**. Data from the miRNA overexpression screen and the CRISPR-Cas9 knockout screen identified two major points of intersection that i) validated both approaches, ii) provided stronger evidence to the identified pathways in promoting resistance, and iii) supported the sophistication of use of two screens to identify novel miRNA-target interactions. Findings from Chapters 2 and 3 identifying such novel miRNA-target interaction partners will be presented in this chapter.

4.1.1 Gene regulation by microRNAs:

MiRNAs function by targeting transcripts of genes via an effector complex composed of the mature miRNA and a core protein, Argonaute (AGO), collectively referred to as the miRNA-induced silencing complex (miRISC). However, when a the transcript arises from a protein coding gene and is negatively regulated by a miRNA, it results in downregulation of the functional product, the protein (7,8). Since proteins are the functional unit of a cell, regulating various normal cellular processes via multiple signaling pathways, alterations in their levels give rise to diseased states such as cancer. In cancerous state of cells, certain proteins are upregulated (oncogenic proteins)

and/or a few others are downregulated (tumor-suppressive proteins) resulting in cellular transformation and progression of cancer. Therefore, understanding the role of miRNAs in regulation of transcripts of such oncogenic and tumor-suppressive proteins, resulting in dysregulation of signaling pathways in the context of cancer is being actively investigated.

4.1.2 MiRNAs mediate gene regulation via non-canonical targeting mechanisms:

Canonically, a miRNA post-transcriptionally regulates a repertoire of transcripts that code for proteins, resulting in mRNA degradation or translational repression. This canonical interaction of miRNA and target transcript is governed by the interaction of the seed sequence of the miRNA (nucleotides 2-7) and the 3′- untranslated region (UTR) region of the target transcript (7,8). Therefore, focus of commonly used bioinformatic tools used to predict the miRNA-target interactome are based on the canonical interaction of miRNA and target 3′-UTR (8,11), excluding the 5′-UTR and coding sequence of the gene.

However, in the light of current research, several non-canonical mechanisms of miRNA mediated transcript targeting have become clear (12). A miRNA can mediate transcriptional or post-transcriptional repression or activation by direct or indirect interaction with its target. For example, some miRNAs have been experimentally proven to post-transcriptionally repress target transcripts via the 5′-UTR, such as the targeting of *RUNX3* by miR-532-5p (13). Indeed, functional studies confirm that miRNA targeting occurs regardless of position of the complementary target sites (14,15). The target region of *let-7* (*Caenorhabditis elegans lin-41*-3′-UTR) was *in vitro* transcribed into the 3′- or 5′-UTR of a reporter construct to show that both reporter mRNAs are equally repressed by *let-7* regardless of position (15). Apart from repression via the 5′-UTR of a target transcript, some miRNAs are also reported to bind to the 5′-UTR recruiting transcription factors that then post-transcriptionally enhance the expression of the transcripts (16).

Figure 4.1. Strategies for the generation of targeting vector to knock-out miRNAs, and tools employed to knock-down miRNAs. (I) Modes of recombination that govern the genomic location of incorporation of the transgene. (A) Homologous recombination (HR) allows a site-specific insertion of the transgene via crossing-over of the specific HR sites between the genomic site and the vector, in the presence of the enzyme recombinase. (B) Random insertion results in incorporation of the transgene at a random site in the genome. (II) Various gene-editing used for the generation of gene-targeting vectors (A) Flp/FRT system: Exogenously added or endogenously expressed Flippase (Flp) recombinase allows site-specific recombination with Flp recombinase target (FRT) sites flanking the transgenic miRNA gene. This results in knocking-out the targeted miRNA. (B) Cre-LoxP system: This system is analogous to the Flp/FRT system. The Cre-recombinase catalyzes site-specific recombination between two-LoxP sites flanking the miRNA of interest, resulting in its excision and miRNA knock-out. (C) Transposition is the mode of transgenic vector incorporation widely used in *D. melanogaster*. This mode of insertion of transgene utilizes multiple mechanisms to insert the transgene at a specific transposon site on the genome. The mechanism here explains a two-vector system. Transposase, the enzyme that facilitates transposition is encoded by the transposase vector, and the miRNA gene to be transposed at the transposon site in the genome is encoded by the second vector, flanked by inverted repeat sequences (IRS) and insertional sites (IS), that mediate the transgene exchange with a random transposon. (D) CRISPR-Cas9 system: Cas9-vector and an sgRNA vector are expressed in cells. The Cas9 endonuclease associates with the expressed sgRNA, which guides Cas9 to a homologous region in the genome to generate a double strand (ds) break. The ds break is repaired in an error-prone manner using Non-Homologous End Joining (NHEJ) or a targeting vector is inserted at the breakpoint via Homologous Recombination to generate either a miRNA knock-out or a knock-in of a miRNA and/or a reporter vector, respectively. (III) Strategies to exogenously or endogenously knock-down a mature miRNA (A-B). (A) Various chemical modifications on small miRNA-complementary oligonucleotides, double stranded or single stranded, have successfully been generated, to sequester functional mature miRNAs and inhibit their function. AntagomiRs are ssRNAs conjugated with cholesterol. Locked-Nucleotide Acid (LNA) are generated via the formation of a 2′, 4′methylene bridge in the ribose resulting in a stable bicyclic nucleotide. 2′-MOE are 2′-O-methoxyethyl phosphorothioate modified oligonucleotides. (B) A miRNA sponge depicted here contains multiple binding sites (6-8) for a specific miRNA in the 3′-UTR of a reporter vector. Sequestration of the miRNA results in negative regulation of the reporter and reduced regulation of the endogenous miRNA targets (Adapted from (8)).

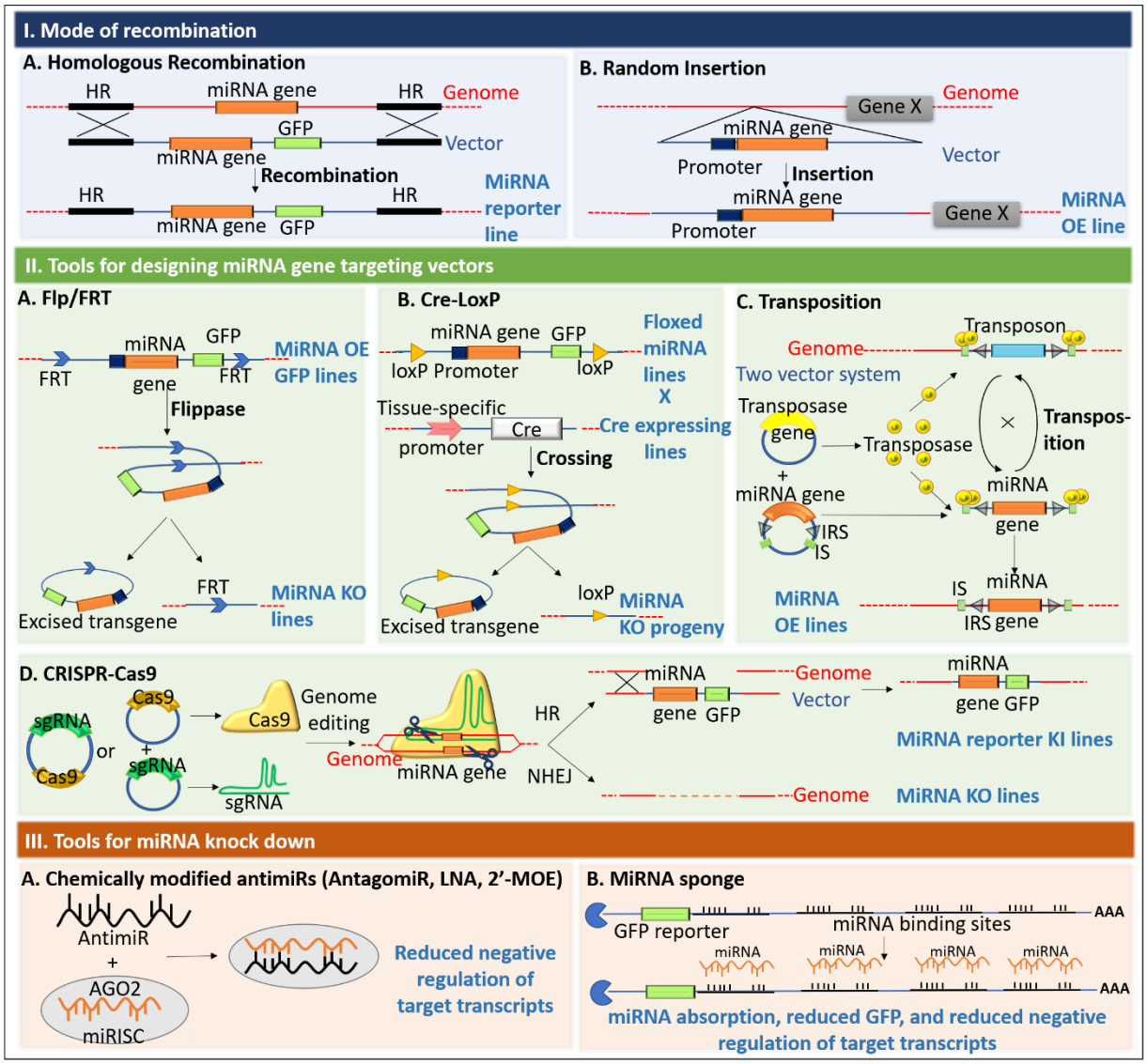
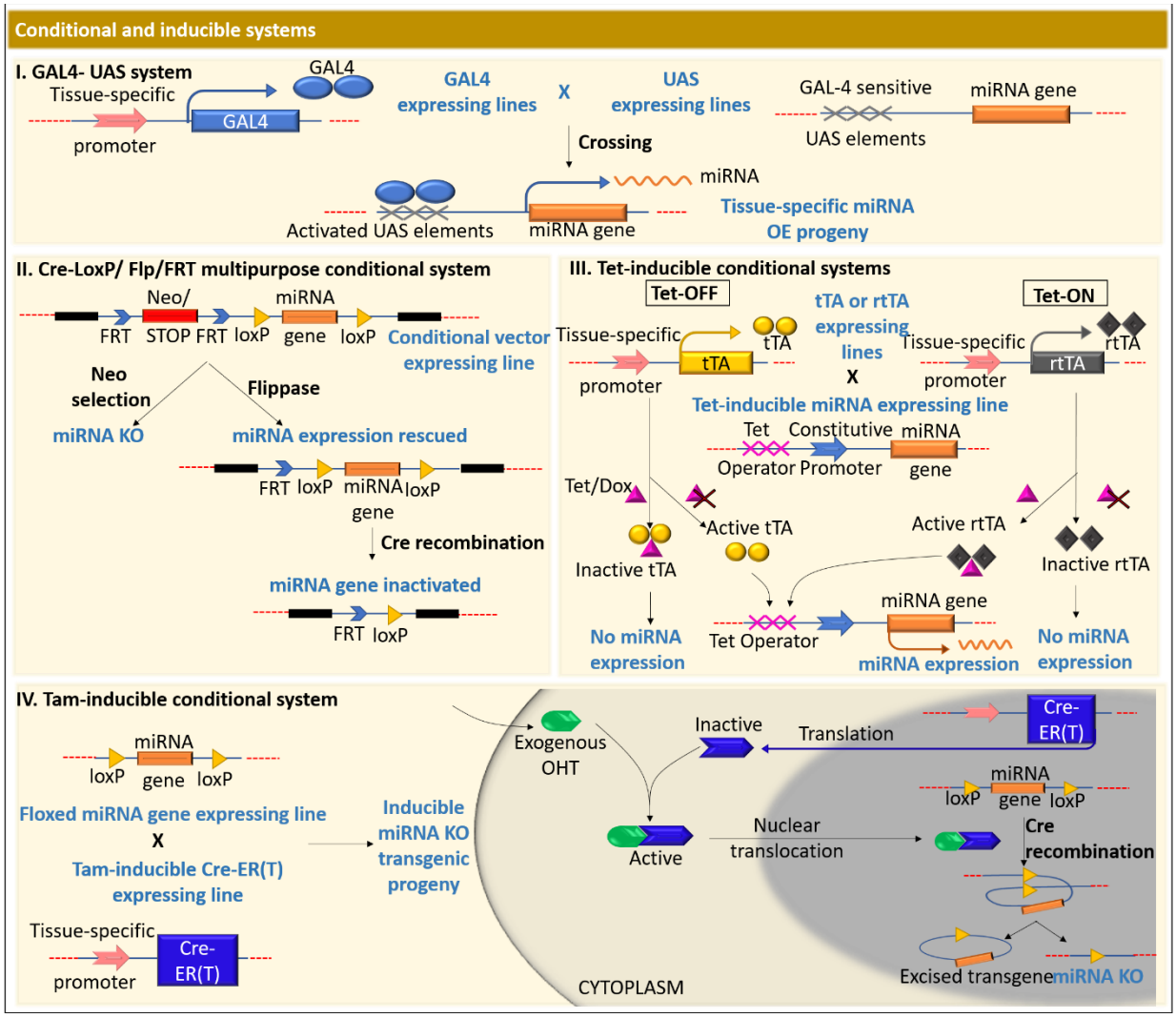

I. Mode of recombination
A. Homologous Recombination
HR
miRNA gene
HR
Genome
GFP
Vector
miRNA gene
Recombination
HR
HR
miRNA gene
GFP
MiRNA reporter line
B. Random Insertion
Gene X
Genome
miRNA gene
Vector
Promoter
Insertion
miRNA gene
Gene X
Promoter
MiRNA OE line
II. Tools for designing miRNA gene targeting vectors
A. Flp/FRT
miRNA
GFP
FRT
gene
FRT
MiRNA OE GFP lines
Flippase
FRT
Excised transgene
MiRNA KO lines
B. Cre-LoxP
miRNA gene
GFP
Floxed miRNA lines
X
loxP Promoter
loxP
Tissue-specific
Cre
Cre expressing lines
promoter
Crossing
loxP
Excised transgene
MiRNA KO progeny
C. Transposition
Transposon
Genome
Two vector system
Transposase
gene
Transposase
Transpos-ition
+
miRNA gene
miRNA
gene
IRS
IS
MiRNA OE lines
IS
miRNA
IRS gene
D. CRISPR-Cas9
Cas9
sgRNA
or
Cas9
+
sgRNA
Cas9
Genome editing
sgRNA
Genome
miRNA gene
HR
NHEJ
miRNA
gene GFP
Genome
Vector
miRNA
gene GFP
MiRNA reporter KI lines
Genome
MiRNA KO lines
III. Tools for miRNA knock down
A. Chemically modified antimiRs (AntagomiR, LNA, 2'-MOE)
AntimiR
+
AGO2
miRISC
Reduced negative regulation of target transcripts
B. MiRNA sponge
GFP reporter
miRNA binding sites
AAA
miRNA
miRNA
miRNA
miRNA
AAA
miRNA absorption, reduced GFP, and reduced negative regulation of target transcripts

Figure 4.2. Conditional and inducible systems. (I) The GAL4/upstream activating sequence (UAS) system (GAL4-UAS) is an inducible system has been utilized in the generation of transgenic flies and zebrafish models. Tissue-specifically expressed GAL4 lines are crossed with a line constitutively expressing the transgene encoded downstream of a UAS element, allowing GAL4 mediated activation of UAS in a tissue-specific manner. Specific binding of GAL4 to UAS element allows the transcription of the miRNA gene, resulting in a tissue-specific overexpression (OE) of the miRNA in the offspring. (II) A combination of Cre-LoxP and Flp/FRT is a powerful tool to generate a multi-purpose conditional and inducible targeting vector. In this case, expression of the Neomycin (Neo)/STOP cassette generates a knock-out first vector, inhibiting the expression of the downstream miRNA gene. However, expression of Flp leads to excision of the STOP cassette through recombination of the two FRT sites, rescuing the miRNA gene expression. This system allows miRNA functional studies first in the absence of miRNA expression, following which the effects of rescuing the miRNA can be evaluated. Finally, the effects of the loss of miRNA can be confirmed upon complete inactivation of the miRNA gene, achieved via expression of Cre. (III) The Tetracycline-inducible systems (Tet-OFF and Tet-ON) have proved to be very versatile in the generation of transgenic model systems. Tet-OFF is mediated via the expression of the Tet transactivator (tTA), whereas the Tet-ON system is dependent on the expression of the reverse tTA (rtTA). Lines expressing tTa or rtTA in a tissue-specific manner are crossed with a transgenic strain expressing the miRNA gene under the control of a constitutive promoter incorporated downstream of a Tetracycline activated element, the Tet Operator (TetO). tTa binds to the TetO in the absence of Tetracycline (Tet) or Doxycycline (Dox), leading to the constitutive expression of the transgenic miRNA gene, while the rtTA remains inactive and unable to bind to TetO in the absence of Tet/Dox inhibiting the expression of the miRNA gene. Upon the addition of Tet/Dox to the Tet-OFF system, tTa binds to Tet/Dox and the miRNA gene expression is turned off, whereas in the case of Tet-ON system, Tet/Dox binds to rtTa enabling it to induce the expression of the miRNA gene via direct interaction with TetO. (IV) Tamoxifen (Tam)-inducible conditional system is an extensively used inducible system in the generation of transgenic model organisms. A strain containing a floxed miRNA gene is crossed to a Tam-inducible Cre-ER(T) expressing line, generating an inducible miRNA knock-out strain. Cre-ER(T) is the Estrogen receptor (ER)-ligand binding domain fused to Cre recombinase, which remains inactive due to sequestration in the cytoplasm. However, upon exogenous addition of hydroxytamoxifen (OHT), the OHT-Cre-ER(T) complex translocates into the nucleus, and actively allows Cre-mediated recombination of the two LoxP sites to occur. The resulting Cre-LoxP recombination knocks-out the miRNA gene from the specific tissue expressing the Tam-inducible Cre-ER(T) vector (Adapted from (8)).


Conditional and inducible systems
I. GAL4- UAS system
Tissue-specific
promoter
GAL4
GAL4
GAL4 expressing lines
X
UAS expressing lines
GAL-4 sensitive
UAS elements
miRNA gene
Crossing
miRNA
Activated UAS elements
miRNA gene
Tissue-specific miRNA OE progeny
II. Cre-LoxP/ Flp/FRT multipurpose conditional system
Neo/
FRT STOP FRT loxP
miRNA
gene loxP
Conditional vector expressing line
Neo selection
miRNA KO
Flippase
miRNA expression rescued
FRT loxP miRNA gene loxP
Cre recombination
miRNA gene inactivated
FRT loxP
III. Tet-inducible conditional systems
Tet-OFF
Tissue-specific
promoter
tTA
tTA
tTA or rtTA expressing lines
X
Tet-ON
Tissue-specific
promoter
rtTA
rtTA
Tet-inducible miRNA expressing line
Tet
Operator
Constitutive
Promoter
miRNA
gene
Tet/Dox
Active tTA
Inactive tTA
Active rtTA
Inactive rtTA
No miRNA expression
Tet Operator
miRNA gene
miRNA expression
No miRNA expression
IV. Tam-inducible conditional system
miRNA
loxP gene loxP
Floxed miRNA gene expressing line
X
Tam-inducible Cre-ER(T) expressing line
Tissue-specific
promoter
Cre-ER(T)
Inducible miRNA KO transgenic progeny
Exogenous OHT
Inactive
Translation
Cre-ER(T)
miRNA
loxP gene loxP
Active
Nuclear translocation
Cre recombination
Excised transgene
miRNA KO
CYTOPLASM

Moreover, another non-canonical miRNA targeting mechanism that is becoming widely accepted is targeting through binding in the coding sequence (CDS) of the transcript. *Let-7* was identified to non-canonically target *DICER*, via three binding sites in its CDS, essentially silencing DICER activity (17). Additionally, miRNAs can directly or indirectly alter transcription of genes via the promoter. Some miRNAs can directly associate with specific motifs on the promoter of a gene, recruiting factors that transcriptionally enhance or downregulate the gene expression (18,19). Alternatively, other miRNAs interact via indirect mechanisms, not targeting the gene directly but instead negatively or positively regulating factors that act as enhancers or repressors of transcription of the gene (20). Such miRNA-transcription factor co-regulatory networks are being actively modelled to identify the various pathways that are indirectly influenced by one miRNA (21,22).

4.1.3 Gene editing by CRISPR-Cas9 system:

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas9 system, derived from prokaryotes has transformed gene manipulation via molecular biology techniques (23). The CRISPR-Cas9 system has proved to be a useful technology to perform in-cell gene editing as well as genome-wide editing (24,25). The CRISPR-Cas9 system can efficiently and precisely knock-in or knock-out a gene of interest (GOI) that aids in accurate evaluation of its function. The basis of gene editing involves the Cas9 enzyme that cleaves double stranded DNA triggering endogenous cellular DNA repair mechanisms (26–29). Specificity is achieved through the use of a small guide RNA (sgRNA) that has complementarity to the GOI, guiding Cas9 to the correct genomic location. Repair is mediated by the cells' error-prone DNA repair mechanism involving non-homologous end-joining (NHEJ). NHEJ causes the incorporation of random insertions or deletions (indels) at the target GOI resulting in loss of gene function (26,27,30).To generate a precise mutation a knock-in strategy is performed by using the CRISPR-Cas9 system. In this case, in addition to Cas9 protein and the targeting small guide RNA (sgRNA), a template that is homologous to a specific genomic locus is introduced. Using a very precise repair mechanism, homology-directed repair (HDR) is performed resulting in incorporation of the template DNA (28,29).

Since the advent of the CRISPR-Cas9 system, several variants of the Cas9 protein have been generated. The guidable catalytically inactive variant, dCas9 is being actively repurposed and used

to transcriptionally or epigenetically regulate genes as well as manipulate and visualize the chromatin (24,25). These approaches entail generating a chimera consisting of dCas9 fused to a protein of interest (i.e. DNA methyltransferase, histone modifier, transcription factor, green fluorescence protein, etc.). Regardless, of the progress in the field, gene-editing by the CRISPR-Cas9 system remains to be a major use of the technology.

4.1.4 CRISPR-Cas9 single guide RNA (sgRNA) mediated gene knock-out:

In the study detailed in Chapter 2, CRISPR-Cas9 mediated gene knock-out was performed to study the effect of loss of genes that mediate erlotinib resistance. For the CRISPR-Cas9 system to function, two components are required, a single guide RNA (sgRNA) and the Cas9 protein (24,25). The Cas9 protein, as described above is required to cleave the target GOI in order to trigger DNA repair mechanism, however, the Cas9 only locates the GOI upon receiving guidance from the sgRNA.

In prokaryotes, the endogenous CRISPR system is guided by a pair of RNAs – the CRISPR RNAs (crRNA) and the trans-activating crRNAs (tracrRNA) (24,25). CrRNA are generated from foreign DNAs such as phage DNA, serving as an adaptive immune mechanism in prokaryotes. The crRNA become incorporated as "spacer sequences" that compose the regions between the Clustered Regularly Interspaced Short Palindromic Repeats (i.e. CRISPR)(31). The crRNA has two defined regions, one that is homologous to the foreign DNA, and the other that base pairs with the tracrRNA (32). Together the crRNA and tracrRNA are required to recruit Cas9 at the target site and cleave it (33).

This prokaryotic system has been taken advantage of for editing specific GOIs in cultured cells and in various *in vivo* systems to study loss of specific GOIs (24,25). Firstly, fusion of the crRNA and tracrRNA to generate a single guide RNA molecule, sgRNA has led to the generation of a very efficient and simplified gene-editing tool (24,25). Secondly, the sgRNA can be custom designed to incorporate 18-24 nucleotides homologous to any specific target GOI but is dependent on the presence of 2-4 nucleotides in the target sites called a protospacer adjacent motif (PAM) (34,35). PAM sequences are critical for target GOI recognition by the sgRNA and DNA cleavage by the Cas9 protein(24,35). Therefore, resultant activity of the CRISPR-Cas9 technology targeting

a GOI is loss of function of the protein product due to incorporation of mutations, and is the system that was adopted for the study in Chapter 2.

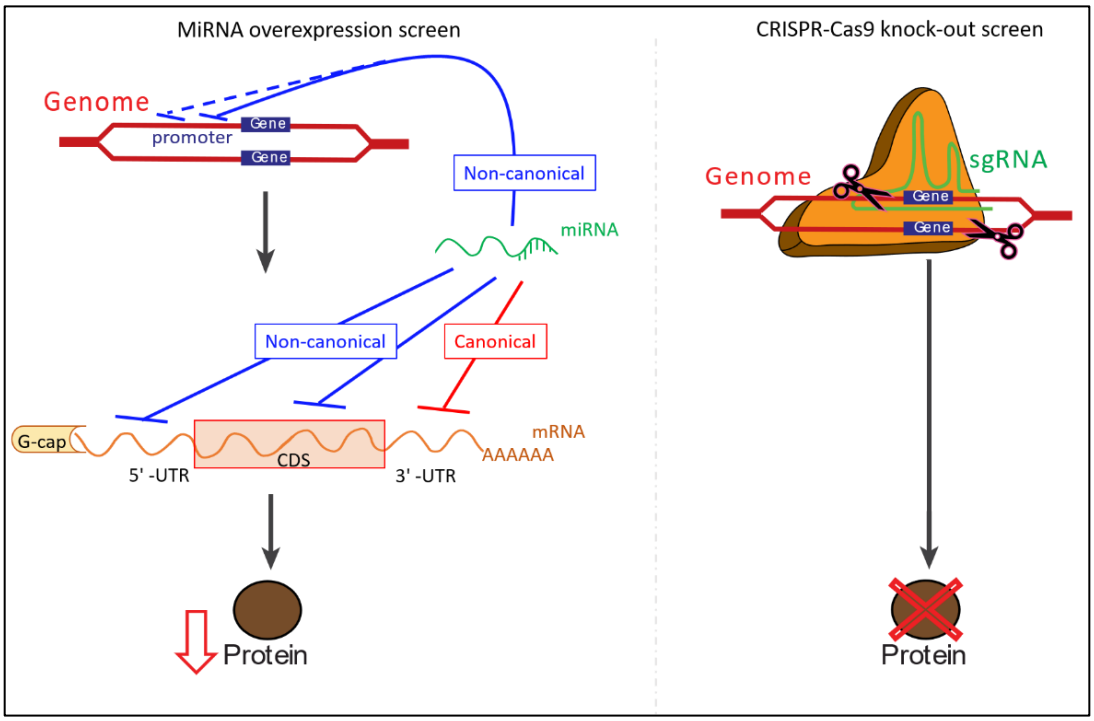


Figure 4.3. Premise of the study. MiRNAs function by canonically or non-canonically targeting genes or transcripts, resulting in downregulation of the protein product. Analogously, CRISPR-Cas9 mediated knock-out of a gene results in incorporation of mutations, and loss of function of the protein product. This study is being conducted to find a correlation between the two screens that identified miRNAs (chapter 3) and proteins (chapter 2) which serve as mediators of erlotinib resistance.

4.1.5 Study design and hypothesis

Both miRNAs and sgRNAs function analogously, i.e. downregulating or abrogating the function of their target, respectively. Therefore, the overall goal of the current study was to validate miRNAs that when overexpressed (identified from the overexpression screen, Chapter 3) result in downregulation of a specific proteins (identified form the knock-out screen, Chapter 2), together functioning as mediators of erlotinib resistance (**Figure 4.3**). The hypothesis is that the miRNA overexpression screen candidates, miR-5693 and miR-4435 target knock-out screen hits, CASP8 and SUV420H2, respectively resulting in the development of erlotinib resistance.

4.2 Methods:

4.2.1 Cell culture:

All cell lines utilized in the study were obtained from American Type Culture Collection (ATCC), cultured under standard conditions and were confirmed to be free of mycoplasma. Cell lines generated during the study were authenticated by ATCC Cell Line Authentication, and were grown in RPMI media supplemented with 10% FBS and 1% Penicillin/Streptomycin cocktail. ECas9 (parental cells, stably expressing Cas9 plasmid) cells were continuously cultured in media containing 1µg/mL Blasticidin, CASP8 knock-out clones 2, 7, and SUV420H2 knock-out clones A, C, E were grown in media containing 100ng/mL Puromycin.

4.2.2 Drug Preparation for *in vitro* studies:

Erlotinib (S7786, Selleck Chemicals) was dissolved in DMSO to prepare 0.4 M stock solutions, which were aliquoted and stored in -80°C. Working concentration of the drug was 200 µM prepared in complete medium and diluted to different concentrations for *in vitro* experiments.

4.2.3 MiRNA overexpression experiments:

To evaluate the effect of miR-5693 and miR-4435 in mediating erlotinib resistance in sensitive NSCLC cells, 6nM miR-5693 (mirVana miRNA mimic, Life Tech, Catalog# 4464066, Assay ID: MC23874) or miR-4435 mimic (mirVana miRNA mimic, Life Tech, Catalog# 4464066, Assay ID: MC20731) were reverse transfected using RNAiMAX.

4.2.4 Gene knockout experiments:

To clone the sgRNA sequences targeting CASP8, the following forward and reverse sequences were utilized: 5′-CACCGGTCATCATCCAGTTTGCATT-3′ (Forward), 5′-AAACAATGCAAACTGGATGATGACC-3′ (Reverse). SUV420H2 was targeted using the following pair of sgRNA sequences (sg2): 5′-CACCGCGGCCCGCTACTTCCAGAGC-3′ (Forward), 5′-AAACGCTCTGGAAGTAGCGGGCCGC-3′ (Reverse) (Integrated DNA Technologies). Each respective oligo pair was annealed and 5′ phosphorylated (T4 Polynucleotide

Kinase kit, M0201S, NEB) as described previously (36). Simultaneously, the CRISPR-Cas9 plasmid, LentiCRISPRv2 (addgene, 52961) plasmid was digested using BsmBI (R0580, NEB), dephosporylated (Antarctic phosphatase, M0289S, NEB) and gel purified using QIAEX II Gel Extraction Kit (20021, Qiagen). The annealed oligos were ligated into the digested and gel purified vector, transformed into Stabl3 bacteria and finally miniprepped, as outlined previously (36). Three µg of pLV-sgCASP8 or pLV-sgSUV420H2 plasmids were linearized and individually forward transfected in $4X10^5$ ECas9 cells (parental cells, stably expressing Cas9 plasmid) using lipofectamine 3000 (L3000015, Thermo Fisher Scientific), following the manufacturer's protocol to finally generate CASP8 knock-out clones 2, 7 and SUV420H2 knock-out clones A, C and E.

4.2.5 Genotyping of mutation:

DNA from ECas9 cells or knock-out clones was isolated using Genomic DNA isolation kit (K1820-01, Thermo Fisher Scientific) following the manufacturer's protocol. CASP8 sequence encompassing the expected mutation was PCR amplified using Q5 high fidelity polymerase (M0491L, NEB) using the following primers: 5′-GTTTACCCTGCAGTTCCTTCT-3′ (Forward), 5′-GTGGATCACGAGGTCAGGAG-3′ (Reverse) and sequenced using the following primers: 5′-CATTTCCCACCACAGGGTCA-3′ (Forward), 5′-AACATGAGCAGCACTTCGGT-3′ (Reverse). SUV420H2 sequence encompassing the expected mutation was PCR amplified similarly using the following primers: 5′- GAGCAGATGGGAGGTGCGGCGACAGT-3′ (Forward), 5′-GAGCTCAGAAGAAAGGAGACAGAT-3′ (Reverse) and sequenced using the following primers: 5′-CCTCTCCTTAGCCTGGTCCT-3′ (Forward), 5′-CAAGGGCTAGGAAGTCAGGG-3′ (Reverse).

4.2.6 RNA isolation and Quantitative real time PCR (qRT-PCR):

Cells ($4X10^5$) were grown in 6 well plates, and total RNA was isolated after 48 or 96 hours, as specified in figure legends, using the miRneasy Kit (217004, Qiagen) according to the manufacturer's instruction. DNase I digestion (79254, Qiagen) was used in each RNA preparation to remove genomic DNA. RNA integrity was evaluated on a 1.5% agarose gel, and total RNA quantified using a nanodrop. cDNA was then synthesized using 1µg of total RNA isolated from cells using MiScript Reverse Transcriptase kit (218161, Qiagen), as indicated by the

manufacturer's protocol. miScript SYBR Green PCR Kit (218073, Qiagen) was utilized as indicated by the manufacturer's protocol, to quantify target gene mRNA expression normalized to the housekeeping, GAPDH via qRT-PCR. Primers used for *CASP8*: 5′-TCATGGACCACAGTAACATGGA-3′ (forward), 5′- AGTGAACTGAGATGTCAGCTCAT-3′ (reverse) (Integrated DNA Technologies), *KMT5C*: 5′- TCGGTTTCCGCACCCATAAG-3′ (forward), 5′-CGGAGGTAGCGATAGACGTG-3′ (reverse) (Integrated DNA Technologies), *GAPDH* (loading control) (QT00079247, Qiagen)

4.2.7 Western Blot:

Cells ($4X10^5$) were grown in 6 well plates, and lysates were isolated at time points as specified in figure legends, using RIPA buffer (Sodium chloride (150 mM), Tris-HCl (pH 8.0, 50mM), N P-40 (1 %), Sodium deoxycholate (0.5 %), SDS (0.1 %), ddH2O (up to 100 mL)) containing 1X protease inhibitor cocktail (PIA32955, Thermo Fisher Scientific). Protein quantification was performed using Pierce BCA Protein Assay kit. Equal amounts of protein lysate were loaded onto a 12% or 4-20% polyacrylamide gel, transferred onto a polyvinylidene difluoride (PVDF) membrane, blocked using LI-COR buffer for 1 hour at room temperature, and incubated overnight in primary antibody at 4 °C. The primary antibody was detected using 1:800 IR 800CW secondary antibody (Li-Cor Biosciences), blots were scanned, and data quantified by using the Odyssey CLx LI-COR imaging system and software (Li-Cor Biosciences). Antibodies used: mouse CASP8 (9746, Cell Signaling), rabbit H4K20me3 (ab9053, abcam), mouse β-ACTIN (3700, Cell Signaling)

4.2.8 Erlotinib dose response

The protocol followed to evaluate erlotinib dose response of various NSCLC cells and cells generated in this study was as per the NCI-60 Cell Five-Dose Screen (NCI-60, DTP (37)). Briefly, Sulforhodamine B colorimetric assay (SRB assay) was performed by exposing cells to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. Post data normalization, as described in figure legends, GI50 erlotinib was calculated from the respective dose curves NCI-60, DTP (37)).

4.2.9 Bioinformatic analyses:

TargetScan – Human(11) was used to predict canonical, and RNA22 v2.0(38) and miRSearch v3.0(39) were used to predict non-canonical targetome of miRs -5693 and -4435. Gene Expression Profiling Interactive Analysis (GEPIA) database (57) (http://gepia.cancer-pku.cn/) was used to evaluate *CASP8* and *KMT5C* in lung adenocarcinoma (LUAD) relative to non-tumorigenic tissue.

4.2.10 Statistical analysis:

All data were analyzed using GraphPad Prism version 7 software (GraphPad Software) and are presented as mean values ± SEM. Student's t-test or one-way ANOVA were the statistical analyses performed, as specified in the figure legends. P-value of < 0.05 was considered significant.

4.3 Results:

4.3.1 *CASP8* and SUV420H2 (*KMT5C*) are low in lung adenocarcinoma (LUAD)

Twenty-five of the predicted thirty-five top hits from the knock-out screen (described in detail in Chapter 2, **Table 2.3**) were identified to be protein-coding genes. Both CASP8 and SUV420H2 have been reported to function as *bona fide* tumor suppressor genes (40–45). Moreover, activation of CASP8 has been positively associated with response to erlotinib (45) and other drugs(46,47). To confirm that *CASP8* and *KMT5C* are present at low levels in lung tumors, as expected for tumor-suppressor genes, we bioinformatically evaluated the levels of *CASP8* and *KMT5C* in LUAD using GEPIA analysis. The results indicate that tumors of LUAD patients express lower levels of both *CASP8* and *KMT5C* relative to normal tissues, suggesting potential tumor-suppressive functions of both gene products in in LUAD (**Figure 4.4 A**). Moreover, prognostic analyses of *CASP8* and *KMT5C* suggested that both genes positively correlate with disease-free survival - LUAD patients with low levels of either gene trended towards a worse prognosis, again indicating likely function of *CASP8* and *KMT5C* as tumor-suppressors in LUAD (**Figure 4.4 B**). Regardless, since both CASP8 (45,48) and SUV420H2 (from results chapter 2) were determined to be mediators of erlotinib resistance in NSCLC, the study is focused on identifying miRNAs from the overexpression screen, especially the top 5 miRNAs (**Table 4.1**) that may have a role in regulating CASP8 and/or SUV420H2 (KMT5C).

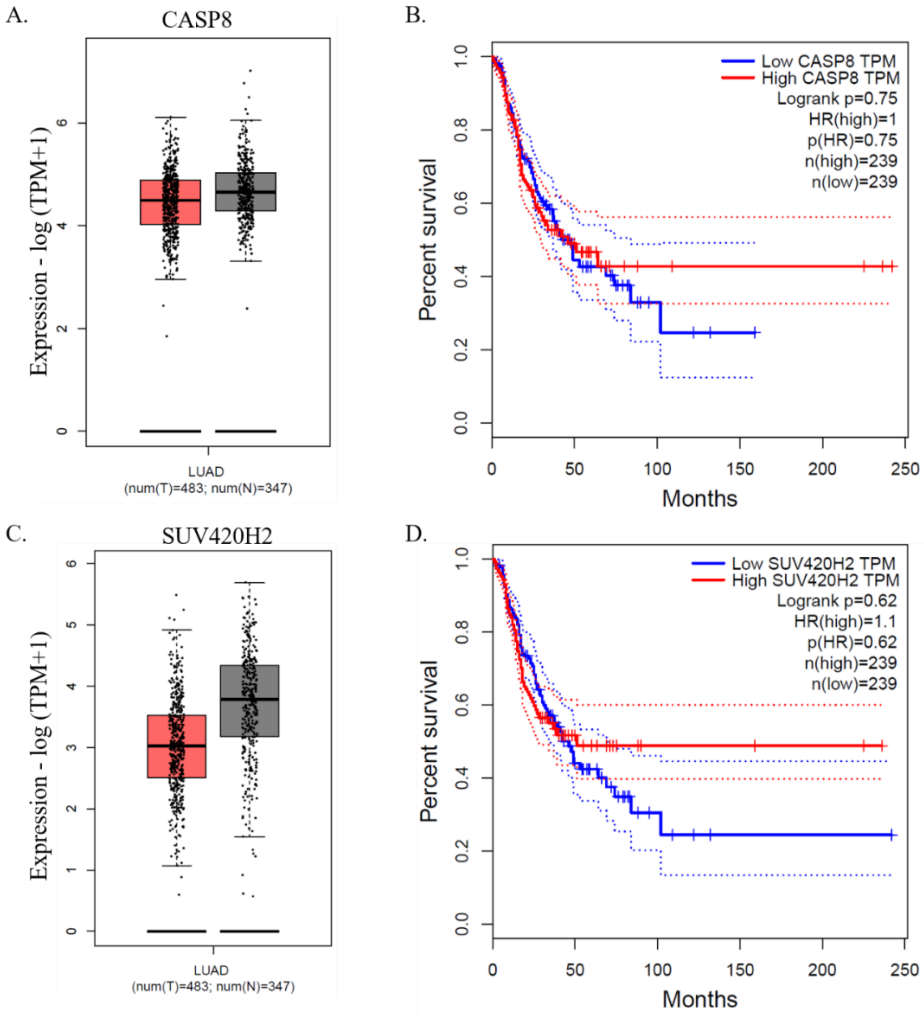


Figure 4.4. Reduced CASP8 and SUV420H2 expressions correlate with poor prognosis in LUAD patients. GEPIA analysis for A) CASP8 and C) *KMT5C* (indicated as SUV420H2 in figures) in normal and tumor samples from LUAD data obtained from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) databases. Disease-free analysis for B) CASP8 and D) *KMT5C* in LUAD patients with high (cutoff = 50%) and low (cutoff = 50%) transcript levels for each gene. TPM= Transcripts per million, T= Tumor, N=Normal.

Table 4.1. Top 5 miRNAs that mediate erlotinib resistance in the sensitive NSCLC cells (EKVX-miR and H322M-pmiR), identified from the overexpression screen. (described in detail in Chapter 3)

Top 5 miRNAs
hsa-miR-5693
hsa-miR-3618
hsa-miR-432-5p
hsa-miR-4435
hsa-miR-588

4.3.2 MiRNAs -5693 and -4435 are predicted to target *CASP8* and *KMT5C*, respectively

MiRNAs -5693 and -4435, candidate miRNAs identified from the overexpression screen belonging to the list of the top five miRNAs (**Table 4.1**) that were validated as mediators of erlotinib resistance, were utilized in this study. Targetome of miR-5693 and miR-4435 predicted using various bioinformatic tools available online, identified genes that were among those identified as contributors to erlotinib resistance when knocked-out (**Figure 4.5** and Chapter 2). Predicted targets of miR-5693 and miR-4435 were retrieved from TargetScan (Human) (11), and were overlapped with the top 25 protein-coding gene transcripts isolated from the knock-out screen (**Figure 4.5 A**). However, since TargetScan (Human) (11) predicts miRNA targetome based on canonical miRNA-target binding, another bioinformatic tool was evaluated that utilizes algorithms to predict non-canonical miRNA binding to its targets, RNA22 v2.0 (**Figure 4.5 B**). Along with canonical binding of miRNAs to targets, RNA22 v2.0 predicts non-canonical binding of miRNAs to target CDS or the 5′-UTR. From **Figure 4.5**, it can be observed that CASP8 was a consistently predicted target of miR-5693. TargetScan identified complementary sequence from miR-5693 in the 3′-UTR of *CASP8*, but RNA22 2.0 identified miR-5693 binding sites in the coding sequence, predicting the potential importance of *CASP8*.

On the other hand, SUV420H2 which was the top hit of the knock-out screen defined in Chapter 2, was identified to be a target of miR-4435 via only non-canonical targeting, predicted by RNA22 v2.0, and not TargetScan. The predicted binding site of miR-4435 lies in the 5′-UTR of the gene encoding SUV420H2, *KMT5C*. This prediction was validated using miRSearch, which also predicted 5′-UTR mediated targeting of *KMT5C* by miR-4435. However, the targeting regions of *KMT5C* predicted by the two algorithms are different (**Figure 4.6**).

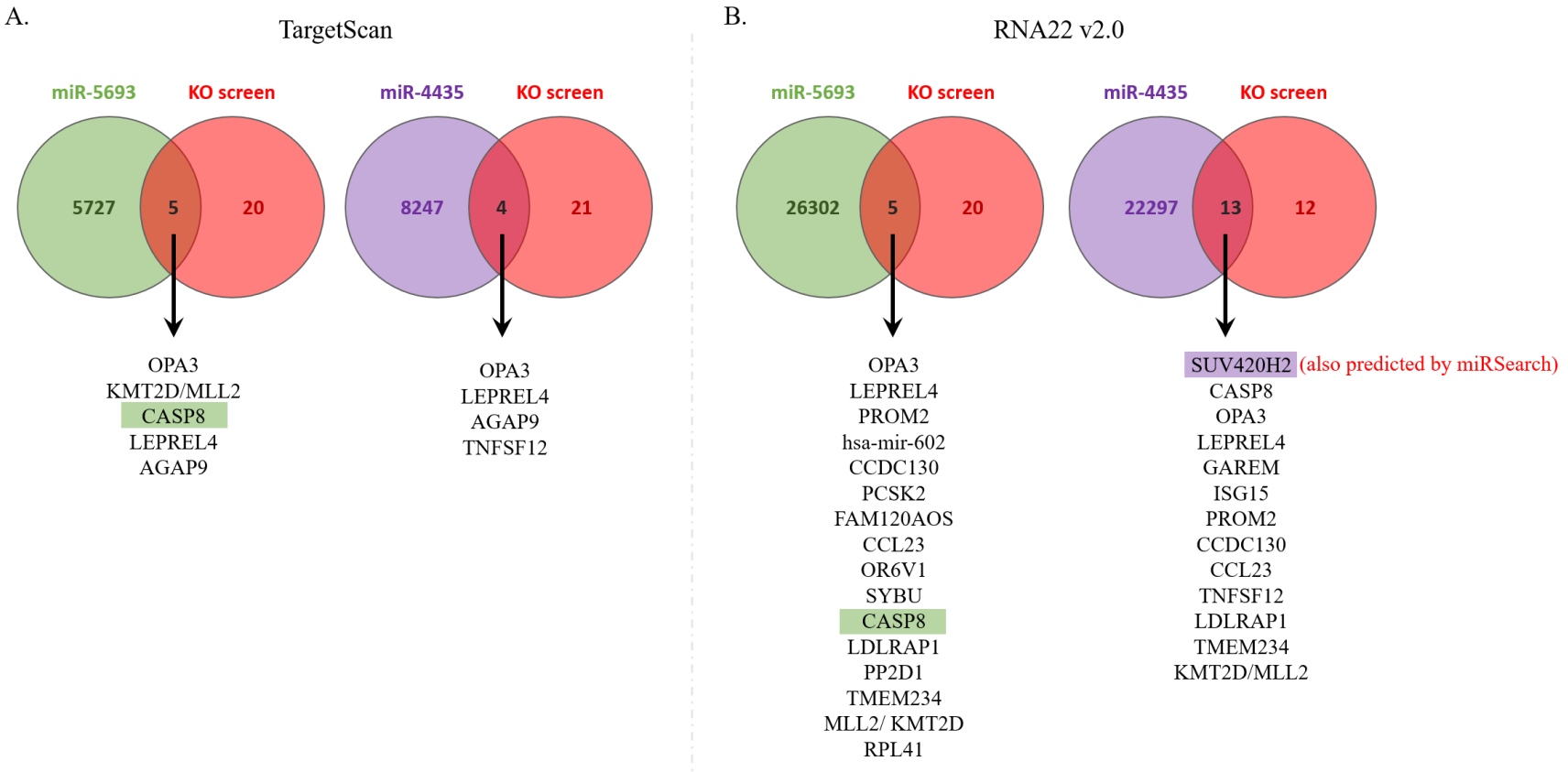


Figure 4.5. Bioinformatically predicted targetome of miRs -5693 and -4435, and venn diagram of targets that overlap with genes identified from the knock-out screen. A) Targets of miR-5693 and miR-4435 predicted by TargetScan (11). TargetScan enlists targetome of miRNAs that are predicted to be canonically regulated. Highlighted in green is CASP8, the target of interest of miR-5693 for this study. B) Targets of miR-5693 and miR-4435 predicted by RNA22 v2.0. RNA22 v2.0 enlists targetome that are predicted to be targeted non-canonically by the miRNAs, i.e. via their CDS or 5′-UTR. Highlighted in purple is SUV420H2, the target of interest of miR-4435 for this study. SUV420H2 is also predicted to be targeted non-canonically by miR-4435 using miRSearch, another miRNA-target prediction tool.

Figure 4.6. Predicted binding sites of miR-4435 on the 5′-UTR of *KMT5C* by two prediction tools, RNA22 v2.0 and miRSearch. Binding site of miR-4435 (b.s.) predicted by RNA22 v2.0 is in orange, b.s. predicted by miRSearch is in blue, and miR-4435 is in green with its seed sequence complementary to the miRSearch b.s. in blue.

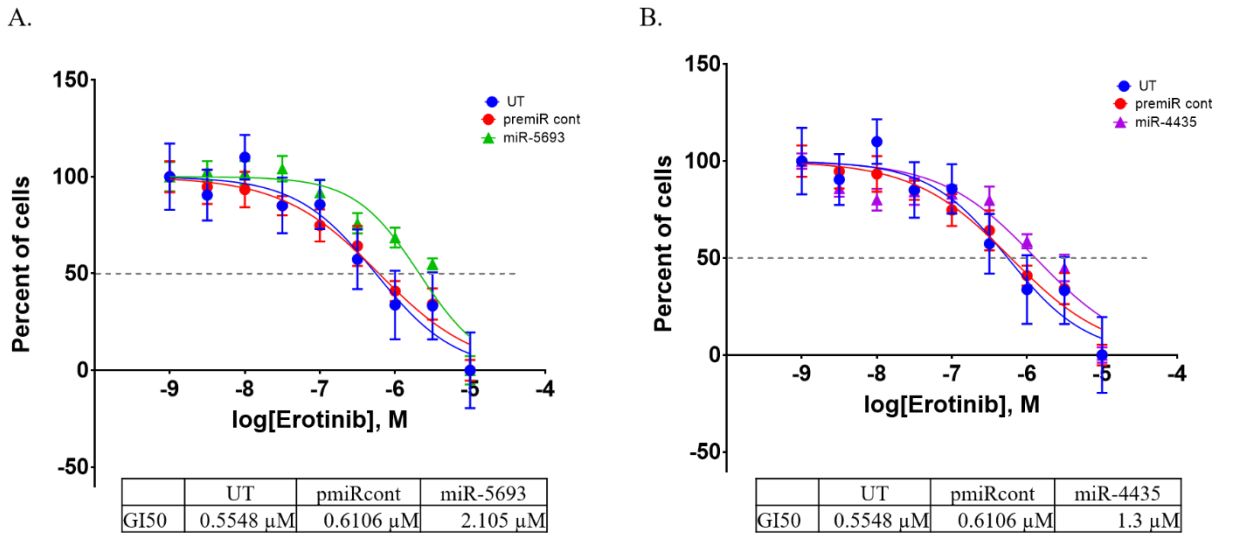
4.3.3 MiR-5693 and miR-4435 enhance erlotinib resistance in the sensitive NSCLC line EKVX

MiRNAs -5693 and -4435, (**Table 4.1**) which consistently validated as mediators of erlotinib resistance, described in Chapter 3, were re-evaluated as mediators of erlotinib resistance in the sensitive NSCLC cell line EKVX. EKVX cells are the parental line from which EKVX-pmiR and EKVX-Cas9 cells were generated and used for the overexpression screen (described in Chapter 3) and the knock-out screen (described in Chapter 2), respectively. Erlotinib dose response post-transfection of miR-5693 or miR-4435 evaluated in EKVX cells confirm that the miRNAs individually enhance erlotinib resistance in the cells (**Figure 4.7**). Similar to the data presented in Chapter 3 for other miRNAs involved in promoting resistance, overexpression of miR-5693 or miR-4435 generated a 2-3 fold increase in resistance.

4.3.4 Loss of CASP8 or SUV420H2 mediates erlotinib resistance in sensitive NSCLC cells

Both CASP8 and SUV420H2 were identified from the knock-out screen as novel mediators of erlotinib resistance. To validate these findings, we first confirmed that sgRNAs targeting the specific genes caused loss of function of the protein. From **Figures 2.3 A**, and **supplementary 2.3 A** (Chapter 2), it can be observed that *KMT5C* was lost in single clones generated post transfection of *KMT5C* targeting sgRNA plasmid. Loss of functional SUV420H2, and its downstream effector, H4K20me3 (**Figures 2.3 B, C**, and **supplementary 2.3 B** (chapter 2)) support the use of

H4K20me3 as a proxy for SUV420H2 function. Following complete loss of the SUV420H2, development of erlotinib resistance can be observed in **Figures 2.3 D and E**, evaluated by erlotinib dose curve generation and proliferation assay.

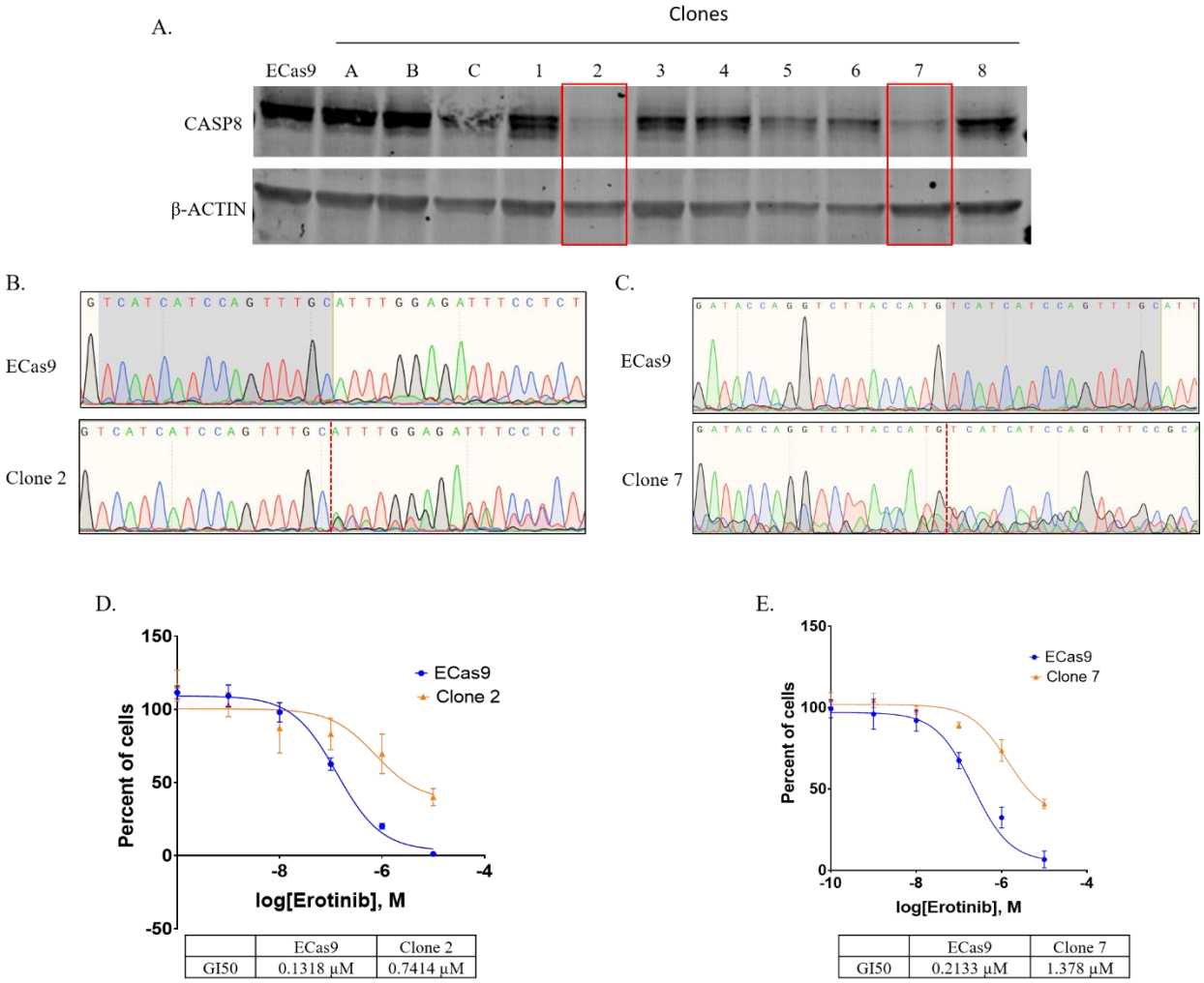


	UT	pmiRcont	miR-5693
GI50	0.5548 µM	0.6106 µM	2.105 µM

	UT	pmiRcont	miR-4435
GI50	0.5548 µM	0.6106 µM	1.3 µM

Figure 4.7. Overexpression of miRNAs, miR-5693 or miR-4435 promote erlotinib resistance in erlotinib sensitive NSCLC cells. Untransfected (UT) EKVX parental cells or EKVX cells reverse transfected with 6nM premiR cont or A) miR-5693 or B) miR-4435 were exposed to increasing concentrations of erlotinib. After 72 hours of exposure to erlotinib (or DMSO control) SRB assays were conducted. For percent of cells calculation, number of cells at the time of addition of erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 erlotinib was calculated from the respective dose curves (as per NCI-60 Cell Five-Dose Screen, NCI-60, DTP (37)).

Similarly, loss of function of CASP8 was also evaluated in single clones generated through stable transfection of *CASP8* targeting sgRNA plasmid (**Figure 4.8 A**). The clones 2 and 7 were further genotyped to confirm *CASP8* mutations mediated by the CRISPR-Cas9 system (**Figure 4.8 B**). Erlotinib resistance mediated by the loss of *CASP8* was validated by erlotinib dose curve generation (**Figure 4.8 C**). *CASP8* mutant clones are 5-fold (clone 2) and 6-fold (clone 7) more resistant to erlotinib.



	ECas9	Clone 2
GI50	0.1318 μM	0.7414 μM

	ECas9	Clone 7
GI50	0.2133 μM	1.378 μM

Figure 4.8. Loss of CASP8 enhances erlotinib resistance in sensitive NSCLC ECas9 cells. A) Western blot image of CASP8 levels in ECas9 cells and clones A-C, 1-8. β-ACTIN was used as a loading control. Western blot was performed using lysates isolated after 48 hours of culturing in complete media. Clones 2 and 7 highlighted in red, have low levels of CASP8, and were taken forward for further studies. Genomic DNA of ECas9 cells or B) clone 2 and C) clone 7 were isolated, the region targeted by CRISPR-Cas9 sgRNA targeting *CASP8* was PCR amplified, purified and sequenced. Representative chromatograms of the wildtype *CASP8* in ECas9 with sgRNA targeting region boxed in grey, and the heterozygous mutations in clones 2 and 7 are represented starting with red dotted lines. Erlotinib dose response via SRB assay was evaluated by exposing ECas9 cells or D) clone 2, and E) clone 7 to varying concentrations of Erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. For percent of cells calculation, number of cells at tz was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 concentration of Erlotinib was calculated from their respective dose curves for each cell line.

4.3.5 miR-5693 and miR-4435 repress CASP8 or SUV420H2, respectively

To determine if miR-5693 and miR-4435 downregulate their predicted targets, CASP8 and SUV420H2, respectively, miRNAs were overexpressed and both transcript and protein levels of the predicted targets were evaluated. Overexpression of miR-5693 did not reduce *CASP8* levels but had modest repressive effects on the CASP8 protein levels (**Figure 4.9 A, C**). MiRNAs can alter their targets through one of two mechanisms, either through translational repression that does not alter the mRNA level or through mRNA destabilization. It is likely that miR-5693 does not result in degradation of the transcript, but may impart translational repression, hence no apparent change in transcript levels, but modest decrease in protein levels. Whereas for SUV420H2 repression mediated by miR-4435, *KMT5C* levels were reduced by approximately 50%, which translated into a modest effect on H4K20me3, measured via western blot (**Figure 4.9 B, D**).

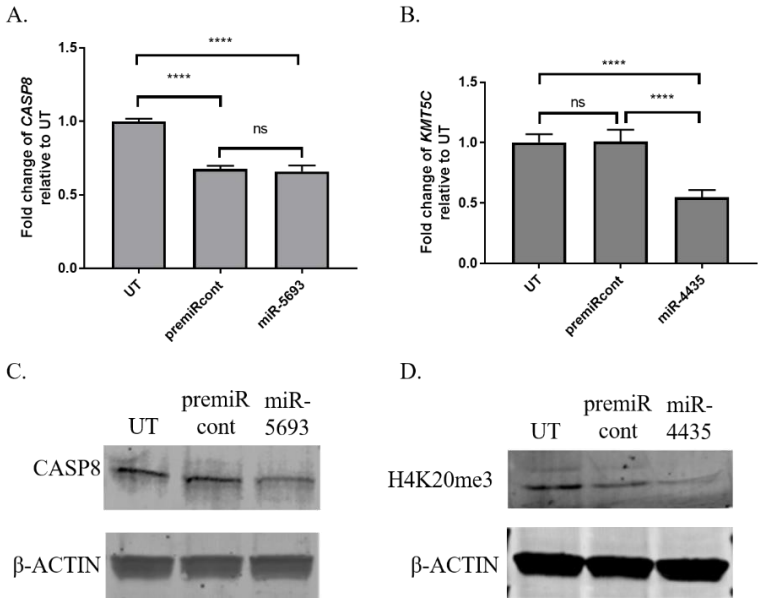


Figure 4.9. MiR-5693 and miR-4435 downregulate CASP8 and SUV420H2 levels, respectively. A) q-RT PCR of *CASP8* levels in untransfected (UT) ECas9 or premiRcont (premiR control) or miR-5693 transfected cells. B) q-RT PCR of *KMT5C* levels in untransfected (UT) ECas9 or premiRcont (premiR control) or miR-4435 transfected cells. *GAPDH* was utilized as the endogenous control for A and B. One-way ANOVA was utilized to evaluate statistical significance of *CASP8* or *KMT5C* transcript levels in A and B, respectively. C) Western blot image of CASP8 levels in UT ECas9 or premiRcont or miR-5693 transfected cells. D) Western blot image of H4K20me3 levels in UT ECas9 or premiRcont or miR-4435 transfected cells. β-ACTIN was utilized as a loading control for C and D. Western blot was performed using lysates isolated 72 hours post-transfection.

4.4 Discussion and future directions:

The findings of this study validate that loss of function of CASP8, identified from the CRISPR-Cas9 knock-out screen in chapter 3 mediates erlotinib resistance in sensitive NSCLC cells (**Figure 4.8**). This is analogous to the loss of SUV420H2, which was the top hit of the CRISPR-Cas9 screen that was validated in Chapter 2. This study also confirmed that miR-5693 and miR-4435, identified amongst the top 5 candidates from the miRNA overexpression screen are *bona fide* mediators of erlotinib resistance (**Figure 4.7**).

However, since the goal of this study was to delineate a correlation between the two screens described previously, an extensive bioinformatic analysis was first conducted to predict if the top five miRNAs targeted any of the 25 protein-coding genes of the knock-out screen. Using the most commonly used bioinformatic tool, TargetScan predicted CASP8 as a putative target of miR-5693. But only one bioinformatic tool identified SUV420H2 as a target of miR-4435, i.e. mirSearch. MiRSearch uses an algorithm that predicts the binding of a miRNA via its seed sequence to the entire sequence of an mRNA, opposed to just the 3′ UTR, thereby predicting non-canonical targeting of miRNAs as well (39,49,50) (**Figure 4.5, 4.6**). This prediction was corroborated by another bioinformatic tool, RNA22 v2.0 (38) that also predicts miRNA-target binding via non-canonical mechanisms throughout the mRNA sequence (**Figure 4.5, 4.6**). Since miR-4435-SUV420H2 or miR-5693-CASP8 have never been reported to function as miRNA-target partners mediating erlotinib resistance together, they were validated experimentally (**Figures 4.7, 4.8 and 4.9**). At this point, it is confirmed that miR-5693 represses CASP8 (shown experimentally via two biological replicates of western, but one of qRT-PCR), and miR-4435 downregulates both *KMT5C* and H4K20me3 (experimentally validated via two biological replicates), but the mechanism of action of the miRNAs are yet to be evaluated.

From the bioinformatic analyses of CASP8 (**Figure 4.5**), it can be observed that both the bioinformatic tools predict miR-5693 binding to CASP8, i.e. via canonical and non-canonical mechanisms. It is therefore likely that miR-5693 mediates its post-transcriptional repression of *CASP8* transcript via either the 3′-UTR (predicted by TargetScan) or the CDS (predicted by RNA22 v2.0), or both. To identify the region of miR-5693 mediated targeting of *CASP8*, the 3′-UTR of *CASP8* can be cloned into a reporter vector (pmiRGLO) and co-transfected with the miRNA to evaluate repression of reporter activity. If no change is observed, it is likely that the

miR-5693 negatively regulates *CASP8* via targeting of its CDS. MiRNAs targeting CDS of a target have been reported to more efficiently inhibit translation of the transcript without degrading it (51,52) (**Figure 4.9**). To test if miR-5693 negatively regulates CASP8 via its CDS, a CDS reporter construct is required to be evaluated with and without miR-5693 transfection.

On the other hand, for miR-4435 mediated downregulation of SUV420H2 evaluation, the 5′-UTR of SUV420H2 will need to be cloned into a pGL3 vector, followed by co-transfection with the miRNA and analysis of luciferase reporter activity. To determine which of the two predicted miR-4435 binding sites on *KMT5C* is involved in repressing *KMT5C* 5′(**Figure 4.6**) mutagenesis studies is currently ongoing. Each of the individual binding sties will need to be mutated to determine the effect of de-repression between the two sites. It is possible that each site on the 5′-UTR may be responsible for some of the repression or that a single site may contain all of the targeting activity. This analysis will validate the predicted functional binding site of the miRNA to the transcript. Additionally, as one miRNA can regulate multiple target transcripts, as suggested by the bioinformatic analyses (**Figure 4.5**), it is likely that miR-5693 and miR-4435 function by targeting other candidates of the knock-out screen as well. Other genes identified through the knock-out screen that were validated, OPA3 and KMT2D (or MLL2) (Chapter 6) are both predicted targets of miRs -5693 and -4435. Therefore, miR-5693 and miR-4435 may likely mediate erlotinib resistance in sensitive NSCLC cells by downregulating multiple targets that were identified through the knock-out screen, a hypothesis that can be pursued in the future.

4.5 References:

1. Qureshi IA, Mehler MF. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. Nat. Rev. Neurosci. 2012. page 528–41.
2. Fox GE, Woese CR. The architecture of 5S rRNA and its relation to function. J Mol Evol . 1975;6:61–76. Available from: http://dx.doi.org/
3. Lerner MR, Steitz JA. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A . 1979;76:5495–9. Available from: http://dx.doi.org/
4. Filipowicz W, Kiss T. Structure and function of nucleolar snRNPs. Mol Biol Rep . 1993;18:149–56. Available from: http://dx.doi.org/

5. Li J, Lin C, Zhang T, Zuo Y, Liu J, Li X, et al. Long noncoding RNA LINC01510 promotes the growth of colorectal cancer cells by modulating MET expression. Cancer Cell Int . BioMed Central; 2018;18:1–12. Available from: https://doi.org/10.1186/s12935-018-0503-5
6. Cech TR, Steitz JA. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. Cell . 2014;157:77–94. Available from: http://doi.org/10.1016/j.cell.2014.03.008
7. Shah MY, Calin GA. MicroRNAs as therapeutic targets in human cancers. Wiley Interdiscip Rev RNA . 2014 [cited 2020 Jan 24];5:537–48. Available from: http://doi.wiley.com/10.1002/wrna.1229
8. Pal AS, Kasinski AL. Animal Models to Study MicroRNA Function. Adv Cancer Res. Academic Press Inc.; 2017. page 53–118.
9. Matsui M, Prakash TP, Corey DR. Argonaute 2-dependent regulation of gene expression by single-stranded miRNA mimics. Mol Ther. Nature Publishing Group; 2016;24:946–55.
10. Thomson DW, Bracken CP, Szubert JM, Goodall GJ. On Measuring miRNAs after Transient Transfection of Mimics or Antisense Inhibitors. PLoS One. 2013;8.
11. No Title . 2015. Available from: http://www.targetscan.org/vert_50/seedmatch.html
12. Cipolla GA. A non-canonical landscape of the microRNA system. Front Genet . 2014;5:337. Available from: http://dx.doi.org/10.3389/fgene.2014.00337
13. Zhang J, Zhou W, Liu Y, Liu T, Li C, Wang L. Oncogenic role of microRNA-532-5p in human colorectal cancer via targeting of the 5'UTR of RUNX3. Oncol Lett. Spandidos Publications; 2018;15:7215–20.
14. Kloosterman WP, The Hubrecht Laboratory Centre for Biomedical Genetics 3584 C T Utrecht The Netherlands, Wienholds E, The Hubrecht Laboratory Centre for Biomedical Genetics 3584 C T Utrecht The Netherlands, Ketting RF, The Hubrecht Laboratory Centre for Biomedical Genetics 3584 C T Utrecht The Netherlands, et al. Substrate requirements for let-7 function in the developing zebrafish embryo. Nucleic Acids Res . 2004;32:6284–91. Available from: https://academic.oup.com/nar/article-pdf/32/21/6284/3846578/gkh968.pdf
15. Lytle JR, Yario TA, Steitz JA. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5′ UTR as in the 3′ UTR. Proc Natl Acad Sci U S A. 2007;104:9667–72.

16. Liu M, Roth A, Yu M, Morris R, Bersani F, Rivera MN, et al. The IGF2 intronic miR-483 selectively enhances transcription from IGF2 fetal promoters and enhances tumorigenesis. Genes Dev. 2013;27:2543–8.
17. Forman JJ, Legesse-Miller A, Coller HA. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc Natl Acad Sci U S A. 2008;105:14879–84.
18. Zhang Y, Fan M, Zhang X, Huang F, Wu K, Zhang J, et al. Cellular microRNAs up-regulate transcription via interaction with promoter TATA-box motifs. RNA. Cold Spring Harbor Laboratory Press; 2014. page 1878–89.
19. Younger ST, Corey DR. Transcriptional gene silencing in mammalian cells by miRNA mimics that target gene promoters. Nucleic Acids Res . 2011 [cited 2020 Jan 25];39:5682–91. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr155
20. Mullany LE, Herrick JS, Wolff RK, Stevens JR, Samowitz W, Slattery ML. MicroRNA-transcription factor interactions and their combined effect on target gene expression in colon cancer cases. Genes Chromosom Cancer. Blackwell Publishing Inc.; 2018;57:192–202.
21. Zhang H-M, Kuang S, Xiong X, Gao T, Liu C, Guo A-Y. Transcription factor and microRNA co-regulatory loops: important regulatory motifs in biological processes and diseases. Brief Bioinform . 2015 [cited 2020 Jan 25];16:45–58. Available from: https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbt085
22. Li R, Chen H, Jiang S, Li W, Li H, Zhang Z, et al. CMTCN: A web tool for investigating cancer-specific microRNA and transcription factor co-regulatory networks. PeerJ. PeerJ Inc.; 2018;2018.
23. Horvath P, Barrangou R. CRISPR/Cas, the immune system of Bacteria and Archaea. Science (80-.). 2010. page 167–70.
24. Adli M. The CRISPR tool kit for genome editing and beyond. Nat. Commun. Nature Publishing Group; 2018.
25. Pickar-Oliver A, Gersbach CA. The next generation of CRISPR–Cas technologies and applications. Nat. Rev. Mol. Cell Biol. Nature Publishing Group; 2019. page 490–507.

26. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science (80-). American Association for the Advancement of Science; 2013;339:819–23.
27. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science (80-). American Association for the Advancement of Science; 2013;339:823–6.
28. Aird EJ, Lovendahl KN, St. Martin A, Harris RS, Gordon WR. Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. Commun Biol. Nature Research; 2018;1.
29. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife. 2014;3:e04766.
30. Guo T, Feng YL, Xiao JJ, Liu Q, Sun XN, Xiang JF, et al. Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. Genome Biol. BioMed Central Ltd.; 2018;19.
31. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol. 2005;60:174–82.
32. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471:602–7.
33. Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P, Siksnys V. crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. RNA Biol. Taylor and Francis Inc.; 2013;10:841–51.
34. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (80-). American Association for the Advancement of Science; 2012;337:816–21.
35. Kleinstiver BP, Prew MS, Tsai SQ, Topkar V V., Nguyen NT, Zheng Z, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. Nature Publishing Group; 2015;523:481–5.
36. LentiCRISPRv2 and lentiGuide-Puro: lentiviral CRISPR/Cas9 and single guide RNA. [cited 2020 Jan 12]; Available from: http://www.genome-engineering.org/gecko/

37. NCI-60 Screening Methodology | NCI-60 Human Tumor Cell Lines Screen | Discovery & Development Services | Developmental Therapeutics Program (DTP) . [cited 2019 Dec 30]. Available from: https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm
38. miRNA Target Visualization . [cited 2020 Jan 25]. Available from: https://cm.jefferson.edu/rna22/Precomputed/OptionController?species=HomoSapiens&type=mRNA&version=MB21E78v2
39. miRSearch - Find microRNAs that target your gene . [cited 2020 Jan 26]. Available from: https://www.exiqon.com/miRSearch?tokenRefresh=true
40. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37:391–400.
41. Yokoyama Y, Matsumoto A, Hieda M, Shinchi Y, Ogihara E, Hamada M, et al. Loss of histone H4K20 trimethylation predicts poor prognosis in breast cancer and is associated with invasive activity. Breast Cancer Res . 2014 [cited 2020 Jan 9];16:R66. Available from: http://breast-cancer-research.biomedcentral.com/articles/10.1186/bcr3681
42. Shinchi Y, Hieda M, Nishioka Y, Matsumoto A, Yokoyama Y, Kimura H, et al. SUV420H2 suppresses breast cancer cell invasion through down regulation of the SH2 domain-containing focal adhesion protein tensin-3. Exp Cell Res . Elsevier; 2015;334:90–9. Available from: http://dx.doi.org/10.1016/j.yexcr.2015.03.010
43. Fulda S, Debatin KM. IFNγ sensitizes for apoptosis by upregulating caspase-8 expression through the Stat1 pathway. Oncogene. Nature Publishing Group; 2002;21:2295–308.
44. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. Immunity. Cell Press; 2018;48:812-830.e14.
45. Orzáez M, Guevara T, Sancho M, Pérez-Payá E. Intrinsic caspase-8 activation mediates sensitization of erlotinib-resistant tumor cells to erlotinib/cell-cycle inhibitors combination treatment. Cell Death Dis. 2012;3:1–9.
46. Okouoyo S, Herzer K, Ucur E, Mattern J, Krammer PH, Debatin K-M, et al. Rescue of death receptor and mitochondrial apoptosis signaling in resistant human NSCLCin vivo. Int J Cancer . 2004 [cited 2020 Jan 9];108:580–7. Available from: http://doi.wiley.com/10.1002/ijc.11585

47. Luo B, Wing H, Subramanian A, Sharifnia T, Okamoto M, Yang X, et al. Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci. 2008;105:20380–20385.
48. Li YT, Qian XJ, Yu Y, Li ZH, Wu RY, Ji J, et al. EGFR tyrosine kinase inhibitors promote pro-caspase-8 dimerization that sensitizes cancer cells to DNA-damaging therapy. Oncotarget . 2015;6:17491–500. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26036637
49. Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol. 2010;18:1139–46.
50. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. Cell Press; 2005. page 15–20.
51. Hausser J, Syed AP, Bilen B, Zavolan M. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. Genome Res. 2013;23:604–15.
52. Zhang K, Zhang X, Cai Z, Zhou J, Cao R, Zhao Y, et al. A novel class of microRNA-recognition elements that function only within open reading frames. Nat Struct Mol Biol . 2018 [cited 2020 Jan 26];25:1019–27. Available from: http://www.nature.com/articles/s41594-018-0136-3

CHAPTER 5. SUMMARY AND FUTURE DIRECTIONS

Chapter Overview

Genome-wide screens conducted to study the roles of specific genes or miRNAs in the context of erlotinib resistance in the previous studies (Chapters 2 and 3), have generated a few candidate genes/miRNAs of interest. However, to dissect true hits from false-positives require validation through meta-analyses, bioinformatic analyses, and *in vitro* experimentations before moving forward into *in vivo* studies. In this chapter, we highlight evidence and preliminary results for a few such candidates that may serve as interesting future directions.

5.1 Loss of SUV420H2 mediates erlotinib resistance:

The top hit identified from the CRISPR-Cas9 knock-out screen, results of which are described in Chapter 2, validated that loss of SUV420H2 causes the development of erlotinib resistance in non-small cell lung cancer cells (NSCLC). It was confirmed that SUV420H2, a histone modifier that represses transcription of genes and maintains heterochromatin via catalyzing trimethylation of Histone H4- Lysine 20 (H4K20), mediates erlotinib resistance by global loss of the H4K20me3 modification. We delineated one mechanism by which loss of SUV420H2 resulted in the upregulation of a long non-coding RNA (lncRNA), *LINC01510* via the loss of its regulatory H4K20me3 modification (**Figures 2.5 and 2.6**). However, since SUV420H2 serves as a genome-wide H4K20me3 facilitator, it can be hypothesized that there are additional genes that are also enhanced due to its' loss. Therefore, the potential directions to explore include: 1) mechanisms by which SUV420H2 performs its function as a regulator of gene expression, and 2) identification of novel downstream effectors of SUV420H2 via which it mediates erlotinib resistance.

5.1.1 Loss of SUV420H2 enhances erlotinib resistance by upregulating MIR4435-2HG

The results of the overexpression screen presented in chapter 3, identified miR-4435 as one of the top miRNAs that promoted erlotinib resistance (**Table 3.1 and Figure 4.7 B**). MiR-4435 has two isoforms, miR-4431-1 and miR-4435-2 that originate from separate genomic loci on chromosome 2 (**Figure 5.1**). Recent reports from the literature support that *MIR4435-2HG* (also annotated as *AK001796, LINC00978*) from which miR-4435-2 originates is an oncogenic lncRNA in multiple

cancers. In lung cancer, *MIR4435-2HG* functions as an enhancer of cell viability and tumorigenicity, by regulating cell cycle progression (1–5). In hepatocellular cancer tissues, *MIR4435-2HG* expression was identified as a poor prognostic predictor, and knockdown of *MIR4435-2HG* in cell lines suppressed cell proliferation and invasiveness (6). Additionally, in NSCLC, prostate, and ovarian cancers *MIR4435-2HG* enhanced cell migration and invasiveness via induction of TGF-β signaling, whereas in gastric cancer cells the phenotype was induced by upregulating Wnt/β-catenin signaling pathway (7–10). In breast cancer cells, it was reported that in addition to development of epithelial-to-mesenchymal (EMT), *MIR4435-2HG* overexpression also enhanced stemness of cancer cells, both of which are hallmarks of aggressive cancers (11,12). Additionally, *LINC00152*, a shorter paralog of *MIR4435-2HG*, that originates from a different genomic locus is also reported to function as a non-coding oncogene, loss of which causes cell cycle arrest in pro-metaphase (13).

In light of these recent findings, it is evident that *MIR4435-2HG* is a *bona fide* oncogenic lncRNA. Additionally, thus far both paralogs, *LINC00152* and *MIR4435-2HG* have been associated with poor responsiveness to certain chemotherapeutic drugs (4,14,15), hence, it is possible that *MIR4435-2HG* has an unidentified role in the development of resistance to targeted therapies such as erlotinib. Based on the findings from the overexpression screen, that miR-4435 promotes erlotinib resistance in NSCLC cells (**Figure 4.7 B**), it is possible that miR-4435 is mediating a part of the function of *MIR4435-2HG*. It can therefore be hypothesized that overexpressing the lncRNA, *MIR4435-2HG* could promote a more dramatic resistant outcome.

Moreover, since loss of SUV420H2 results in a heightened erlotinib resistant phenotype (**Figure 2.3**), we evaluated if a link exists between the MIR4435-2HG and SUV420H2 that can be explored. Bioinformatic analysis using the H4K20me3 ChIP-seq results from the work published by Nelson and colleagues, GSE59316 (16) was conducted. The results revealed that the genes for the two isoforms of miR-4435, -1 and -2 are located downstream of H4K20me3 modifications. MIR4435-2HG hosts a H4K20me3 modification in its gene body, which is upstream of the *miR-4435-2* precursor sequence (**Figure 5.1B**). Additionally, the *miR-4435-1* precursor that originates from a distinct genomic locus also has a H4K20me3 modification upstream of its sequence. (**Figure 5.1A**).

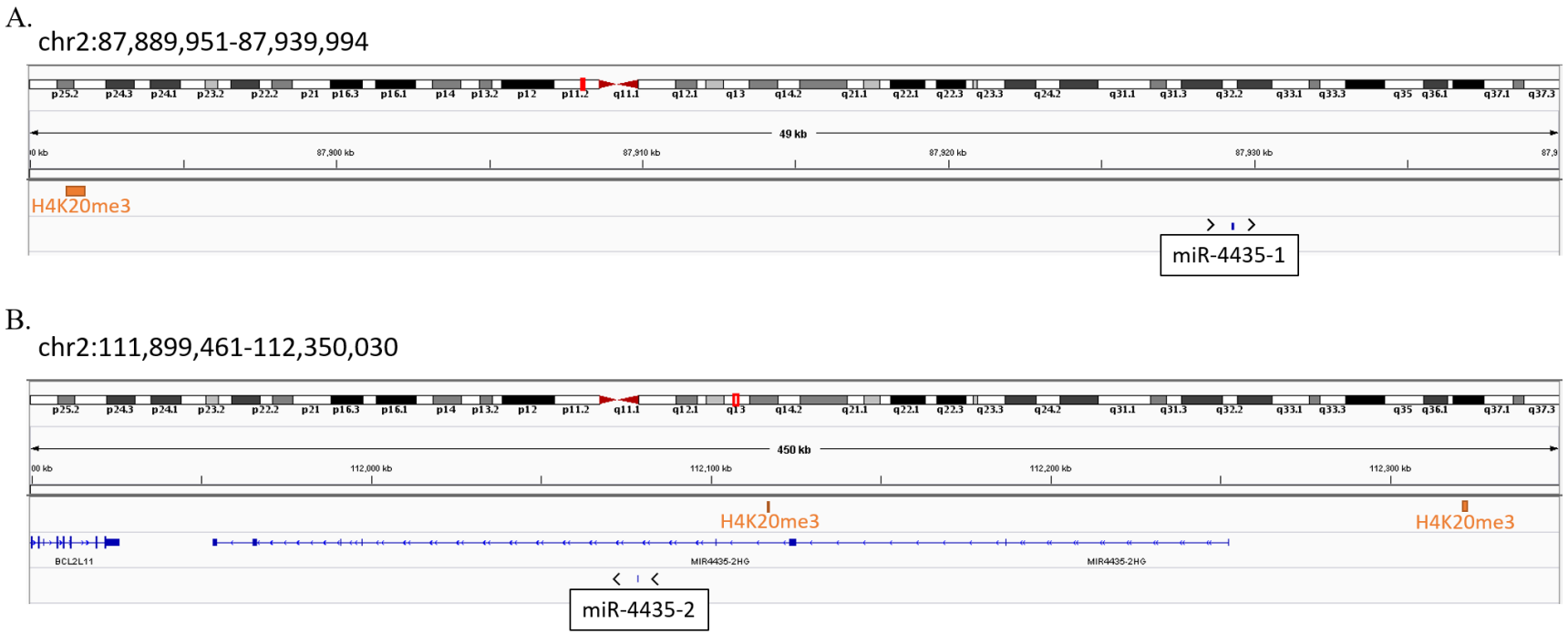


Figure 5.1. Genomic loci of miR-4431-1 and miR-4435-2, and the associated H4K20me3 marks upstream. Integrated Genome Viewer (IGV) was used to visualize GSE59316 (16). The genomic locus of the miRNA isoform A) miR-4435-1 and B) miR-4435-2 are depicted in the above diagrams. H4K20me3 modifications are indicated in orange, and the arrows around the genomic loci of each miRNA (black) depict direction of transcription

From ChIP-qPCR results it was established that H4K20me3 modifications found on either the gene body or upstream of the gene can both negatively regulate gene expression (**Figure 2.5, 2.6, Supplementary Figure 2.6**). It is therefore possible that in erlotinib sensitive cells with functional SUV420H2 that *MIR4435-2HG* is downregulated due to high H4K20me3 levels either on the gene body or upstream of the gene. But in the case of erlotinib resistant cells, it can be hypothesized that with reduced SUV420H2 activity, H4K20me3 mediated repression of *MIR4435-2HG* is released resulting in accumulation of oncogenic *MIR4435-2HG*. To evaluate this potential negative correlation between *KMT5C* (the transcript of SUV420H2) and *MIR4435-2HG*, preliminary bioinformatic analysis was performed with NSCLC samples (The Cancer Genome Atlas (TCGA) dataset) using GEPIA (17) (**Figure 5.2A**). The result suggests a modest but a significant negative correlation (Pearson's correlation = 0.-12) between *KMT5C* and *MIR4435-2HG*. Additionally, to confirm oncogenicity of *MIR4435-2HG*, overall survival analysis was conducted for NSCLC samples (TCGA) using TANRIC (18), suggesting an inverse correlation between *MIR4435-2HG* and NSCLC prognosis (**Figure 5.2B**).

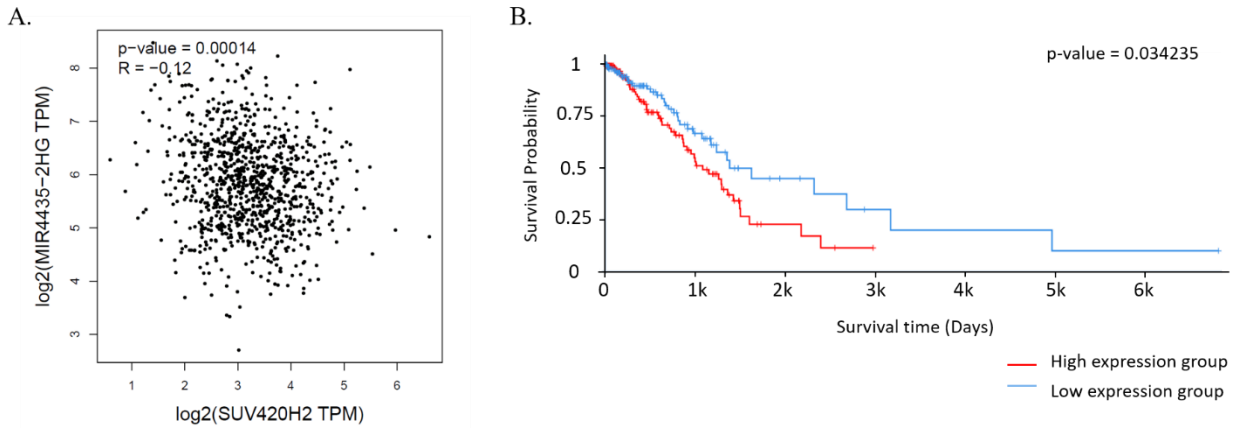


Figure 5.2. *KMT5C* negatively correlates with *MIR4435-2HG*, a poor prognostic marker of NSCLC. A) GEPIA database (http://gepia.cancer-pku.cn/)(17) was used to evaluate a correlation between *MIR4435-2HG* and *KMT5C* (depicted as SUV420H2 in results) in NSCLC patient samples. B) *MIR4435-2HG* (ENSG00000172965.10) associated with overall survival in NSCLC samples obtained from TCGA, and analyzed by TANRIC (18).

To further evaluate the importance of SUV420H2-MIR4435-2HG interaction in erlotinib resistance, a few approaches can be employed. Firstly, evaluating the expression of *MIR4435-2HG* in SUV420H2 knock-out cell lines, which are erlotinib resistant may indicate a possible role for *MIR4435-2HG* as a mediator of erlotinib resistance if highly expressed. Knock-down and knock-

out experiments of *MIR4435-2HG* in the SUV420H2 knock-out background followed by assessing erlotinib response will identify how much of the SUV420H2 knock-out effects are mediated via *MIR4435-2HG*. Simultaneously, overexpressing *MIR4435-2HG* in erlotinib sensitive cells, followed by evaluating erlotinib response will determine if *MIR4435-2HG* truly functions as a mediator of erlotinib resistance. Next, ChIP-qPCR for the predicted locus in normal cells versus SUV420H2 knock-out cells, or in SUV420H2 overexpressing cells, will delineate H4K20me3 mediated regulation of *MIR4435-2HG* via SUV420H2. If *MIR4435-2HG* validates as a *bona fide* mediator of erlotinib resistance, RNA-seq can be performed to identify genes that are significantly up or downregulated following modulation of *MIR4435-2HG*. Further validations will be required to identify downstream effectors of *MIR4435-2HG* as a mediator of erlotinib resistance.

5.1.2 SUV420H2 selectively and dynamically regulates lncRNA genes, and maintains heterochromatin

SUV420H2 is necessary for the maintenance of heterochromatin regions of the genome, such as the telomeres, and as a repressor of oncogenes that if spuriously activated can cause cancers (19–21). Other important regions of the genome that are regulated by the H4K20me3 mark include repetitive regions of the genome (19,21,22), and at least one other report, apart from our results from Chapter 2, shows that the H4K20me3 modification regulates multiple lncRNAs (23).

LINC01510 was identified as an oncogenic lncRNA (reported in Chapter 2) that partially imparts resistance in erlotinib sensitive cells expressing low levels of SUV420H2 (**Figures 2.7, 2.8**). To study *LINC01510* and SUV420H2 dynamics, cell lines were generated that conditionally express SUV420H2 following doxycycline (DOX) treatment (**Figures 2.4A, B**). When cells were grown continually in the presence of DOX for four days, the H4K20me3 modification increased relative to control treated cells (**Figure 5.3A**). It should be noted that H4K20me3 levels increase in both parental Calu6 and PBS treated clone 2 cells over time likely due to low levels of SUV420H2 and proliferation, and thus increased cell number over time. However, it cannot be ruled out that other enzymes catalyzing H4K20me3 may be involved in these cell lines (24,25). Regardless, clone 2 cells exposed to dox had a further increase in H4K20me3 at all time points, indicative of SUV420H2 induction following dox treatment.

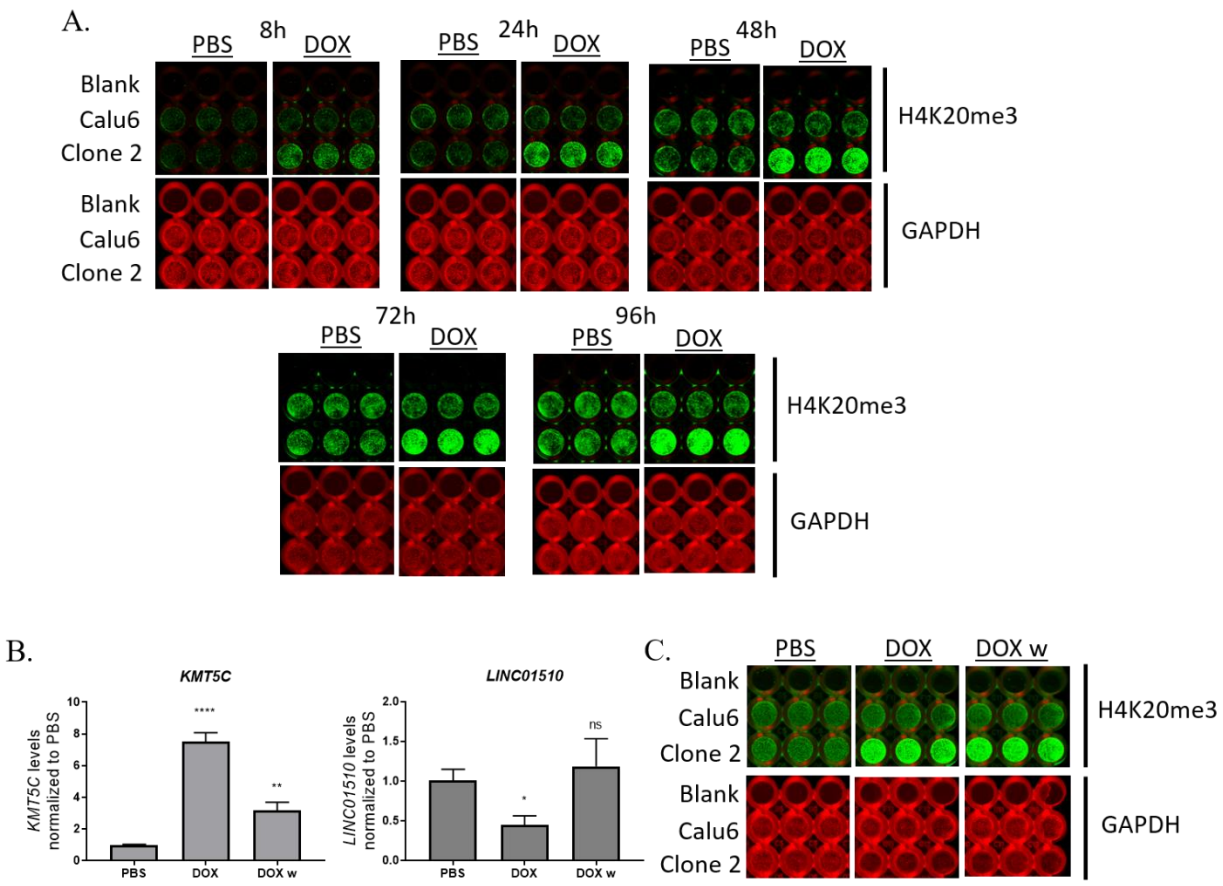


Figure 5.3. Altered SUV420H2 levels maintain overall H4K20me3 mark, but dynamically regulate *LINC01510* levels. A) H4K20me3 levels evaluated by in-cell western (ICW) in Calu6 cells and Calu6 derived clone 2 cells that stably express a DOX-inducible SUV420H2. Cells were grown in complete media and re-plated in media containing PBS or DOX and not changed throughout the assay. ICW was performed at 8, 24, 48, 72, 96 hours after seeding. GAPDH serves as the endogenous control. B) *KMT5C* or *LINC01510* levels evaluated by qRT-PCR. Both KMT5C and LINC01510 are normalized to GAPDH. Cells were grown in PBS or DOX containing media for one week, and re-plated in either PBS containing media, DOX containing media or complete media (DOX withdrawn media, DOX w) for 24 hours. C) H4K20me3 levels evaluated by ICW in Calu6 or clone 2 cells. Treatments are as in B. GAPDH is used as the endogenous control.

To evaluate the dynamics of the H4K20me3 modification cells were grown in their respective PBS or DOX media for a week, and as expected expressed high *KMT5C* levels and low *LINC01510* levels in the presence of DOX (**Figure 5.3B**). However, following the withdrawal of DOX for 24 hours, while the H4K20me3 modification was still robust (**Figure 5.3C**) *LINC01510* expression was restored to wildtype levels (**Figure 5.3B**). This suggests that certain regions of the genome

that are required to be maintained in a repressed state, such as the constitutive heterochromatic regions may be regulated more tightly than others. It is likely that SUV420H2 has a higher affinity for constitutive heterochromatic regions harboring high levels of H4K20me3 mark and is sequestered there to lock-down the chromatin. Thus, when SUV420H2 levels are reduced the soluble fraction of SUV420H2 may become specifically re-localized to the heterochromatin over localization at repressed genes. The dynamics of SUV420H2 and H4K20me3 in cancer and development is largely unknown and is an exciting and active area of investigation in the laboratory, based on these preliminary findings.

The findings from **Figure 5.3** suggest that SUV420H2 is potentially required to maintain the heterochromatin, and since loss of SUV420H2 causes genomic instability and chromatin aberrations (21,26–30); it is therefore possible that the two phenotypes are linked and working concurrently. Thus, it can be hypothesized that the lack of maintenance of heterochromatin via loss of SUV420H2 mediates the observed erlotinib resistance in NSCLC cells, causing massive genomic instability and impaired DNA damage repair mechanism (**Figure 2.9**).

While SUV420H2 is imperative for maintenance of heterochromatin, the lncRNA *LINC01510* appears to be a less critical gene. Exogenous increase of *LINC01510* clearly identified a pro-survival role for *LINC01510* via MET induction, in the case of erlotinib resistance (**Figure 2.8**). However, in the context of absence of erlotinib, it is possible that other lncRNAs are more critical and *LINC01510* is dispensable. As reported by Kurup and group (23), the H4K20me3 modification is important for the regulation of multiple lncRNAs. Therefore, to better understand the normal cellular function of SUV420H2, and its' role in the development of cancer via regulation of the heterochromatin and lncRNAs simultaneously, it is critical to understand SUV420H2 dynamics at the chromatin level.

In order to identify genes that are dynamically regulated by SUV420H2 via the H4K20me3 modification, cells that stably express the DOX inducible SUV420H2 plasmid can be used for the following experiments. Firstly, cells that are unexposed to DOX (i.e. PBS exposed) will serve as control, following which chromatin will be immunoprecipitated (ChIP) using the H4K20me3 primary antibody (as described in Chapter 2) and the associated DNAs will be sequenced (i.e. ChIP-seq) following DOX exposure. The ChIP-seq can be performed at various time points

following DOX treatment to achieve a comprehensive picture of SUV420H2 dynamics. From **Figure 5.3A** it can be observed that SUV420H2 induction leads to accumulation of the H4K20me3 modification on the chromatin within eight hours, which can serve as a good starting point. Other time points can include twenty-four hours, seven days and one month post DOX exposure. Simultaneously, RNA-seq can be performed at the aforementioned time points from both PBS and DOX exposed cells. Finally, overlapping the data from the ChIP- and the RNA- seqs will result in the determination of SUV420H2 regulated genes and their transcripts. DNA sequences that are enriched in DOX relative to PBS exposed cells and their corresponding RNA transcripts that are specifically enriched in PBS relative to DOX cultured cells will be further evaluated. Finally, transcripts that were enriched in the initial phase followed by those that stably remained upregulated over time in PBS treated cells can be identified as early phase transcripts or passengers and late-phase or oncogenic transcripts, respectively. The late-phase transcripts, especially those of lncRNAs will serve as hits for further evaluation as oncogenic lncRNAs that are negatively regulated by SUV420H2. A similar but converse experiment can be performed in EKVX (high SUV420H2 expressing cells) using a DOX-inducible shRNA targeted to SUV420H2. Dynamic changes in the H4K20me3 modification and changes to the transcriptome following time-dependent downregulation of SUV420H2 will underscore the dynamics involved in SUV420H2 loss or reduction at these critical genomic regions.

5.1.3 SUV420H2 and lncRNAs as regulators of heterochromatin

This brief section describes a hypothesis based on a reported phenomenon that SUV420H2 interacts with lncRNAs at the chromatin, to facilitate the formation of heterochromatin (31–33). SUV420H2 can therefore be envisioned as not only a regulator of lncRNAs, as previously described, but that SUV420H2 also requires lncRNAs for its function.

To identify the lncRNAs that are involved in heterochromatin formation by direct interaction with SUV420H2, firstly a specific SUV420H2 antibody is required to be generated. SUV420H2 antibodies from five different vendors failed to identify altered SUV420H2 levels in SUV420H2 knock-out clones relative to parental cells in our laboratory (**Figure 2.3A-C, Supplementary Figure 2.3A, B**). Therefore, to conduct the experiments highlighted below, it is imperative to have a specific antibody that recognizes SUV420H2.

For this study, chromatin-associated RNA immunoprecipitation and sequencing (CARIP-seq) can be a used to delineate SUV420H2-lncRNA dynamics occurring at the chromatin (23). For a preliminary experiment, CARIP-seq results of SUV420H2 knock-out (clone c) cells compared to that of ECas9 cells (parental cells) will identify differentially associated lncRNAs. The lncRNAs critical for SUV420H2 to function in the maintenance of heterochromatin are expected to be severely lost in CARIP-seq results of clone c cells relative to ECas9 cells. The results of this study will identify a cohort of lncRNAs that are dysregulated upon loss of SUV420H2, and are critical for heterochromatin maintenance by interaction with the chromatin via SUV420H2.

The study can be further expanded by using cells that stably express the DOX inducible SUV420H2 plasmid to evaluate if heterochromatin formation in cells is a dynamically regulated process. CARIP-seq can be performed post DOX exposure to induce the activity of SUV420H2 in clone 2 cells at various time points and results compared to respective PBS control cells. The proposed starting time points based on **Figure 5.3A** is eight hours, followed by additional time points such as twenty-four hours, seven days and one month of DOX exposure. The enriched lncRNAs in DOX treated sequencing results relative to that of PBS treated cells will yield a group of lncRNAs that serve as: 1) initiators of heterochromatin formation if enriched at early time points, and 2) stabilizers of heterochromatin if the lncRNAs remain enriched at later time points. The identified lncRNAs will require further molecular and *in vitro* evaluation to determine their role as regulators of the heterochromatin via SUV420H2.

5.2 Additional validated targets from the knock-out screen

SUV420H2 and CASP8 were identified from the knock-out screen as key mediators of erlotinib resistance. In addition to those, a few other genes identified in the screen (**Table 2.3**), that were reported to function as tumor suppressor genes or miRNAs, were partially validated here as drivers of erlotinib resistance.

The protein Outer Mitochondrial Membrane Lipid Metabolism Regulator, OPA3, to the best of our knowledge has been reported to function as a tumor suppressor in only one study (34). The findings show that loss of OPA3 enhanced EMT in TGF-β induced retinal pigment epithelial ARPE-19 cells (34). Since the normal function of OPA3 is to regulate mitochondrial dynamics, a key organelle in drug response (35,36), in addition to its function as a negative regulator of EMT,

it surfaced as one of our candidate genes. Another protein that was selected as the second candidate gene for validation was KMT2D/MLL2, an epigenetic factor i.e. histone-lysine N-methyltransferase 2D. In addition to the reported tumor suppressive role of KMT2D/MLL2 in multiple cancers (37–39), KMT2D/MLL2 was positively associated with sensitivity to chemotherapeutic drugs (other than erlotinib) in lung cancer cells (40).

As for candidate miRNAs, miR-148a was reported to function as a tumor suppressor in one study (41). It was determined that silencing miR-148a in cancer associated fibroblasts results in increased motility of cells via activation of the WNT signaling pathway (41). MiR-512-1 on the other hand has been reported to function as a tumor-suppressive miRNA. In lung cancer, miR-512-1 suppress various hallmarks of cancer, such as cell proliferation, migration, and invasion, and is also reported to induce cell death (42). In breast, and head and neck cancers, miR-512-1 causes suppression of telomerase activity and shortening of telomeres, and loss of miR-512 is associated with aggressive breast cancer (43,44).

In addition to the above miRNAs, the levels of miR-602 and miR-648 were evaluated in the erlotinib sensitive cell NSCLC line, EKVX and a very resistant cell line Calu6 identified from the study conducted by National Cancer Institute, Developmental Therapeutics Program (NCI-60, DTP (45)) (**Figure 5.4**). It was observed that both miRNAs were lower in Calu6 cells relative to EKVX, analogous to the results of the screen, i.e. loss of the miRNAs -602 and -648 result in driving erlotinib resistance, therefore were selected as candidates.

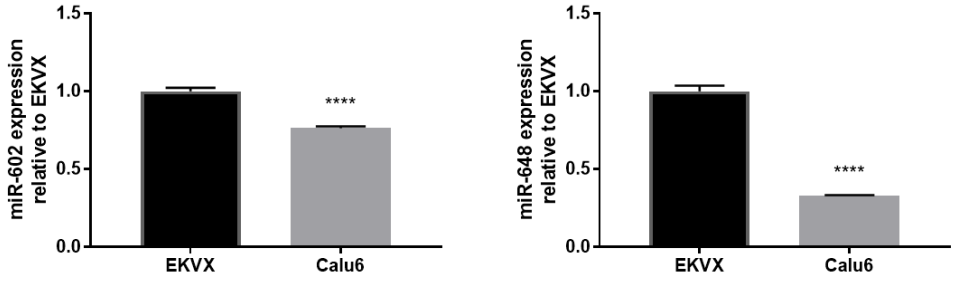


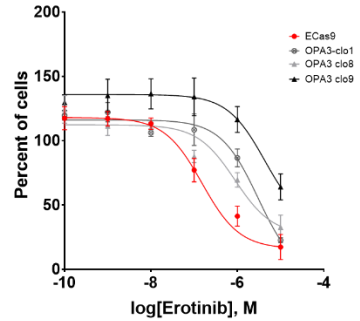
Figure 5.4. MiR-602 and miR-648 levels are low in Calu6 cells relative to EKVX cells. MiR-602 and miR-648 (mature miRNAs) levels were evaluated using qRT-PCR in EKVX and Calu6 cells. GAPDH is utilized as the endogenous control.

The above identified candidate genes (OPA3, KMT2D, miR-148a, miR-512-1, miR-602 and miR-648) from the knock-out screen were mutated using the CRISPR-Cas9 system in the cell line used to conduct the screen, ECas9 cells (generated from EKVX cells). Single clones were generated and validated for their response to erlotinib. The preliminary results suggest that at least two out of the four selected clones for each candidate displayed erlotinib resistance (**Figure 5.5**).

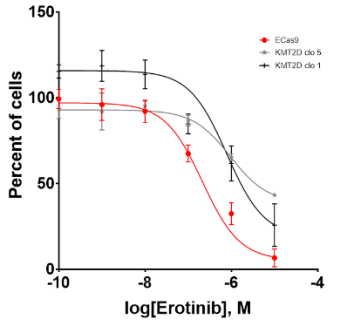
To further validate that candidate gene/s in the above single clone/s were specifically mutated by the CRISPR-Cas9 system, the locus will need to be genotyped. Following which, western blot analyses for OPA3 and KMT2D, and qRT-PCR evaluation of the miRNAs of interest will confirm loss of function of the candidates. The candidate genes that validate for loss of function, in turn mediating the development of erlotinib resistance will be taken forward for further characterization. A panel of cells will be used to identify if a negative correlation exists between the expression of the candidate and GI50 values of the panel. For genes/miRNAs that follow the trend, ectopic overexpressed will be conducted in erlotinib resistant cell lines with low expression levels of the candidate. These cells are expected to become re-sensitized, at least partially to erlotinib. For candidate genes/miRNAs that do re-sensitize resistant NSCLC cells to erltotinib additional molecular characterization to delineate the mechanism by which the candidates alter erlotinib response will be conducted.

5.3 Closing remarks:

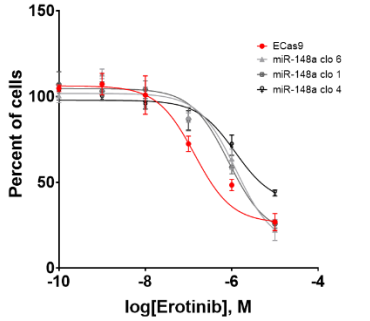
Epigenetics was first acknowledged to play a role in development in 1975 by two simultaneous studies identifying that DNA methylation can turn genes on or off (46,47). This finding resulted in an explosion of studies identifying novel DNA and histone modulators, especially in the context of development and cancer (48–51). Currently, epigenetics in cancer is regarded as a hallmark of cancer (52) and is an actively investigated field. While in the field of microRNAs (miRNAs), in 1993 the first miRNA *lin-4* was identified. *lin-4* functions in the normal development of *Caenorhabditis elegans* (53). Since then research in the field of miRNAs has galvanized towards delineating miRNAs that contribute to the hallmarks of cancer (54). Certain miRNAs can also fall into a class of epigenetic factors if their levels are regulated via epigenetic mechanisms, or if the miRNAs play a role in modulating epigenetic enzymes (55). Such miRNAs have also been reported to be involved in carcinogenesis (54,55).



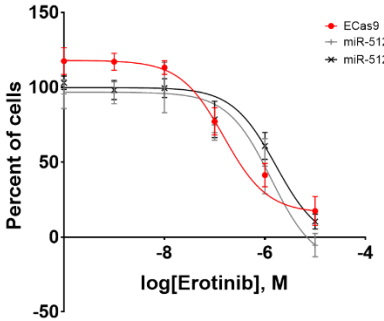
	ECas9	OPA3 clo8	OPA3 clo9	OPA3-clo1
GI50 (M)	1.521e-007	8.785e-007	4.34e-006	3.034e-006



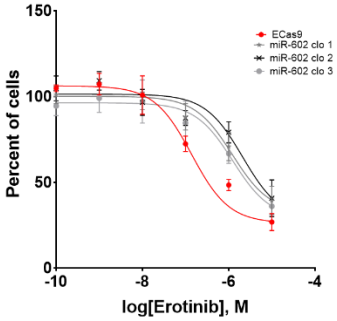
	ECas9	KMT2D clo 5	KMT2D clo 1
GI50 (M)	2.133e-007	9.493e-007	7.824e-007



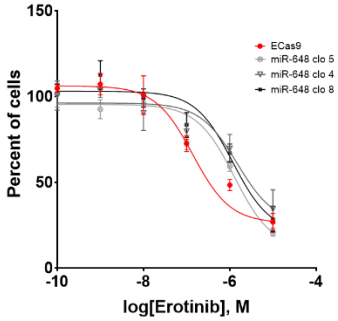
	ECas9	miR-148a clo 6	miR-148a clo 1	miR-148a clo 4
GI50 (M)	1.388e-007	1.26e-006	8.491e-007	1.198e-006



	ECas9	miR-512 clo 3	miR-512 clo5
GI50 (M)	1.521e-007	1.331e-006	1.651e-006



	ECas9	miR-602 clo 1	miR-602 clo 2	miR-602 clo 3
GI50 (M)	1.388e-007	1.215e-006	2.111e-006	1.244e-006



	ECas9	miR-648 clo 4	miR-648 clo 8	miR-648 clo 5
GI50 (M)	1.388e-007	1.497e-006	1.176e-006	1.276e-006

Figure 5.5. Single clones generated after stably expressing CRISPR-Cas9 plasmid targeting the specific candidate genes, generate erlotinib resistant cells. Ten thousand ECas9 parental cells or single clones generated after stably expressing CRISPR-Cas9 that targets a specific candidate gene were plated. Erlotinib dose response via SRB assay was evaluated by exposing cells to varying concentrations of Erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. For percent of cells calculation, number of cells at the time of addition of Erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 erlotinib was calculated from the respective dose curves (as per the NCI-60 Cell Five-Dose Screen, NCI-60, DTP (45)).

The two screens conducted during the course of this study, the miRNA overexpression screen (Chapter 3) and the CRISPR-Cas9 knock-out screen (Chapter 2) were aimed at identifying novel epigenetic mediators of erlotinib. The miRNA, miR-4435 and the epigenetic factor, SUV420H2 were individually identified as critical mediators of erlotinib resistance, and were shown to collaborate to impart resistance (in Chapter 4). However, because the function of SUV420H2 was only first reported in 2004 (19), and was not validated as a contributor to carcinogenesis until 2005 (56); currently there is an immense knowledge gaps regarding the biological role of SUV420H2. Consequently, the various hypothesis and future directions described in section 5.1 will aid in delineation novel mechanisms by which SUV420H2 functions as a tumor suppressor and a mediator of erlotinib resistance. Additionally, as described in chapter 4, miR-4435 a putative SUV420H2 targeting miRNA was independently identified as a mediator or erlotinib resistance, which further validates the critical role for SUV420H2 in promoting erlotinib resistance.

The novelty of conducting the two screens described in the study was two-fold:

1) It not only aided in identifying miRNAs that when overexpressed impart erlotinib resistance, but also identified certain miRNAs from the knock-out screen that can act as potential inhibitors of resistance if overexpressed, described in section 5.2. Currently, as miRNAs are being investigated as therapeutic molecules, the miRNAs identified via the knock-out screen have the potential to be further evaluated as "erlotinib resistance blockers".

2) Using the CRISPR-Cas9 system, genes were first mutated, and mutants that truly developed erlotinib resistance were identified. This is potentially the reason why loss SUV420H2 was identified in this study as causal of erlotinib resistance, despite its function. As can be observed from the data presented in chapter 2, loss of SUV420H2 upregulates multiple oncogenes (including MYC, BMP4, OCT4, and SFTPC, which have also been tested), consequently possibly resulting in the development of a heterogenous population. If the screen was conducted via conventional approaches, wherein cells were first exposed to TKIs followed by sequencing to identify altered genes, SUV420H2 may have remained undiscovered; whereas the oncogenes that accumulated as a result of loss of SUV420H2 may have been identified as mediators of resistance. In conclusion, the screens conducted in the study have undeniably resulted in the discovery of novel epigenetic mediators of erlotinib resistance in NSCLC.

5.4 References:

1. Yang Q, Xu E, Dai J, Liu B, Han Z, Wu J, et al. A novel long noncoding RNA AK001796 acts as an oncogene and is involved in cell growth inhibition by resveratrol in lung cancer. Toxicol Appl Pharmacol. Academic Press Inc.; 2015;285:79–88.
2. Qian H, Chen L, Huang J, Wang X, Ma S, Cui F, et al. The lncRNA MIR4435-2HG promotes lung cancer progression by activating β-catenin signalling. J Mol Med. Springer Verlag; 2018;96:753–64.
3. Ling L, Huang J, Ma S, Luo K, Qian H, Cui F, et al. The lncRNA MIR4435-2HG promotes lung cancer progression by activating β-catenin signalling. J Mol Med. Journal of Molecular Medicine; 2018;96:753–64.
4. Bian Z, Zhang J, Li M, Feng Y, Yao S, Song M, et al. Long non-coding RNA LINC00152 promotes cell proliferation, metastasis, and confers 5-FU resistance in colorectal cancer by inhibiting miR-139-5p. Oncogenesis. Nature Publishing Group; 2017;6.
5. Li X, Ren Y, Zuo T. Long noncoding RNA LINC00978 promotes cell proliferation and invasion in non-small cell lung cancer by inhibiting miR-6754-5p. Mol Med Rep. Spandidos Publications; 2018;18:4725–32.
6. Han Q-L, Chen B-T, Zhang K-J, Xia S-T, Zhong W-W, Zhao Z-M. The long non-coding RNA AK001796 contributes to poor prognosis and tumor progression in hepatocellular carcinoma. 2014.
7. Wang H, Wu M, Lu Y, He K, Cai X, Yu X, et al. LncRNA MIR4435-2HG targets desmoplakin and promotes growth and metastasis of gastric cancer by activating Wnt/�-catenin signaling. Aging (Albany NY) [Internet]. 2019 [cited 2020 Jan 28];11. Available from: http://www.aging-us.com/article/102164/text?_escaped_fragment_=
8. Zhang H, Meng H, Huang X, Tong W, Liang X, Li J, et al. lncRNA MIR4435-2HG promotes cancer cell migration and invasion in prostate carcinoma by upregulating TGF-β1. Oncol Lett [Internet]. 2019 [cited 2020 Jan 28];18:4016–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31516603
9. Gong J, Xu X, Zhang X, Zhou Y. LncRNA MIR4435-2HG is a potential early diagnostic marker for ovarian carcinoma. Acta Biochim Biophys Sin (Shanghai). Oxford University Press; 2019;51:953–9.

10. Yang M, He X, Huang X, Wang J, He Y, Wei L. LncRNA MIR4435-2HG-mediated upregulation of TGF-β1 promotes migration and proliferation of nonsmall cell lung cancer cells. Environ Toxicol. John Wiley and Sons Inc.; 2019;
11. Deng LL, Chi YY, Liu L, Huang NS, Wang L, Wu J. LINC00978 predicts poor prognosis in breast cancer patients. Sci Rep. Nature Publishing Group; 2016;6.
12. Kumar M, Sinnott DeVaux R, Shen JJ, Davis SP, Dinger ME, Mattick JS, et al. Abstract 1598: LncRNA AK001796 as a therapeutic target in aggressive breast cancers. American Association for Cancer Research (AACR); 2016. page 1598–1598.
13. Nötzold L, Frank L, Gandhi M, Polycarpou-Schwarz M, Groß M, Gunkel M, et al. The long non-coding RNA LINC00152 is essential for cell cycle progression through mitosis in HeLa cells. Sci Rep. Nature Publishing Group; 2017;7.
14. Yue B, Cai D, Liu C, Fang C, Yan D. Linc00152 functions as a competing endogenous RNA to confer oxaliplatin resistance and holds prognostic values in colon cancer. Mol Ther. Nature Publishing Group; 2016;24:2064–77.
15. Liu B, Pan CF, Ma T, Wang J, Yao GL, Wei K, et al. Long non-coding RNA AK001796 contributes to cisplatin resistance of non-small cell lung cancer. Mol Med Rep. Spandidos Publications; 2017;16:4107–12.
16. Nelson DM, Jaber-Hijazi F, Cole JJ, Robertson NA, Pawlikowski JS, Norris KT, et al. Mapping H4K20me3 onto the chromatin landscape of senescent cells indicates a function in control of cell senescence and tumor suppression through preservation of genetic and epigenetic stability. 2016;
17. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res [Internet]. 2017 [cited 2020 Jan 12];45:W98–102. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkx247
18. Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, et al. TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. Cancer Res. American Association for Cancer Research Inc.; 2015;75:3728–37.
19. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. 2004;18:1251–62.

20. Marión RM, Schotta G, Ortega S, Blasco MA. Suv4-20h abrogation enhances telomere elongation during reprogramming and confers a higher tumorigenic potential to iPs cells. PLoS One. 2011;6:e25680.
21. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callén E, et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev. Cold Spring Harbor Laboratory Press; 2008;22:2048–61.
22. Kidder BL, Hu G, Cui K, Zhao K. SMYD5 regulates H4K20me3-marked heterochromatin to safeguard ES cell self-renewal and prevent spurious differentiation. Epigenetics and Chromatin. BioMed Central; 2017;10:1–20.
23. Kurup JT, Kidder BL. Identification of H4K20me3 and H3K4me3-associated RNAs using CARIP-Seq expands the transcriptional and epigenetic networks of embryonic stem cells Downloaded from H4K20me3 and H3K4me3-associated RNAs in ES cells. [cited 2020 Jan 28]; Available from: http://www.jbc.org/cgi/doi/10.1074/jbc.RA118.004974
24. Dai L, Ye S, Li H-W, Chen D-F, Wang H-L, Jia S-N, et al. SETD4 Regulates Cell Quiescence and Catalyzes the Trimethylation of H4K20 during Diapause Formation in Artemia. Mol Cell Biol. American Society for Microbiology; 2017;37.
25. Ye S, Ding Y-F, Jia W-H, Liu X-L, Feng J-Y, Zhu Q, et al. Tumor Biology and Immunology SET Domain-Containing Protein 4 Epigenetically Controls Breast Cancer Stem Cell Quiescence. 2019 [cited 2020 Jan 10]; Available from: http://cancerres.aacrjournals.org/
26. Bromberg KD, Mitchell TRH, Upadhyay AK, Jakob CG, Jhala MA, Comess KM, et al. The SUV4-20 inhibitor A-196 verifies a role for epigenetics in genomic integrity. Nat Chem Biol. 2017;13:317–24.
27. Hahn M, Dambacher S, Dulev S, Kuznetsova AY, Eck S, Wörz S, et al. Suv4-20h2 mediates chromatin compaction and is important for cohesion recruitment to heterochromatin. Genes Dev. 2013;27:859–72.
28. Jørgensen S, Schotta G, Sørensen CS. Histone H4 Lysine 20 methylation: Key player in epigenetic regulation of genomic integrity. Nucleic Acids Res. 2013;41:2797–806.
29. Kovaríková AS, Legartová S, Krejcí J, Bártová E. H3K9me3 and H4K20me3 represent the epigenetic landscape for 53BP1 binding to DNA lesions. Aging (Albany NY). Impact Journals LLC; 2018;10:2585–605.

30. Sanders SL, Portoso M, Mata J, Bähler J, Allshire RC, Kouzarides T. Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell. 2004;119:603–14.
31. Zhao Z, Dammert MA, Grummt I, Bierhoff H. LncRNA-Induced Nucleosome Repositioning Reinforces Transcriptional Repression of rRNA Genes upon Hypotonic Stress. Cell Rep. Elsevier B.V.; 2016;14:1876–82.
32. Bierhoff H, Dammert MA, Brocks D, Dambacher S, Schotta G, Grummt I. Quiescence-Induced LncRNAs Trigger H4K20 Trimethylation and Transcriptional Silencing. Mol Cell. Cell Press; 2014;54:675–82.
33. Park J, Lee H, Han N, Kwak S, Lee H-T, Kim J-H, et al. Long non-coding RNA ChRO1 facilitates ATRX/DAXX-dependent H3.3 deposition for transcription-associated heterochromatin reorganization. Nucleic Acids Res [Internet]. 2018 [cited 2020 Jan 28];46:11759–75. Available from: http://crispr.mit.edu/
34. Ryu SW, Yoon J, Yim N, Choi K, Choi C. Downregulation of OPA3 Is Responsible for Transforming Growth Factor-β-Induced Mitochondrial Elongation and F-Actin Rearrangement in Retinal Pigment Epithelial ARPE-19 Cells. PLoS One. 2013;8:1–9.
35. Okon IS, Coughlan KA, Zhang M, Wang Q, Zou M-H. Gefitinib-mediated ROS instigates mitochondrial dysfunction and drug resistance in lung cancer cells. 2015 [cited 2020 Jan 29]; Available from: http://www.jbc.org/cgi/doi/10.1074/jbc.M114.631580
36. Deng J, Shimamura T, Perera S, Carlson NE, Cai D, Shapiro GI, et al. Proapoptotic BH3-Only BCL-2 Family Protein BIM Connects Death Signaling from Epidermal Growth Factor Receptor Inhibition to the Mitochondrion. Cancer Res [Internet]. 2007 [cited 2020 Jan 29];67:11867–75. Available from: http://cancerres.aacrjournals.org/
37. Hillman RT, Celestino J, Terranova C, Beird HC, Gumbs C, Little L, et al. KMT2D/MLL2 inactivation is associated with recurrence in adult-type granulosa cell tumors of the ovary. Nat Commun. Nature Publishing Group; 2018;9.
38. Kantidakis T, Saponaro M, Mitter R, Horswell S, Kranz A, Boeing S, et al. Mutation of cancer driver MLL2 results in transcription stress and genome instability. Genes Dev. Cold Spring Harbor Laboratory Press; 2016;30:408–20.

39. Ardeshir-Larijani F, Bhateja P, Lipka MB, Sharma N, Fu P, Dowlati A. KMT2D Mutation Is Associated With Poor Prognosis in Non–Small-Cell Lung Cancer. Clin Lung Cancer [Internet]. Elsevier Inc.; 2018;19:e489–501. Available from: https://doi.org/10.1016/j.cllc.2018.03.005
40. CtRP v2. Cancer Therapeutics Response Portal https://portals.broadinstitute.org/ctrp.v2.1/ [Internet]. [cited 2020 Jan 9]. Available from: https://portals.broadinstitute.org/ctrp.v2.1/
41. Aprelikova O, Palla J, Hibler B, Yu X, Greer YE, Yi M, et al. Silencing of miR-148a in cancer-associated fibroblasts results in WNT10B-mediated stimulation of tumor cell motility. Oncogene. 2013;32:3246–53.
42. Cao B, Tan S, Tang H, Chen Y, Shu P. MiR-512-5p suppresses proliferation, migration and invasion, and induces apoptosis in non-small cell lung cancer cells by targeting ETS1. Mol Med Rep. Spandidos Publications; 2019;49:3604–14.
43. Dinami R, Buemi V, Sestito R, Zappone A, Ciani Y, Mano M, et al. Epigenetic silencing of miR-296 and miR-512 ensures hTERT dependent apoptosis protection and telomere maintenance in basal-type breast cancer cells. Oncotarget. Impact Journals LLC; 2017;8:95674–91.
44. Li J, Lei H, Xu Y, Tao Z. miR-512-5p Suppresses Tumor Growth by Targeting hTERT in Telomerase Positive Head and Neck Squamous Cell Carcinoma In Vitro and In Vivo. Fan G-C, editor. PLoS One [Internet]. 2015 [cited 2020 Jan 29];10:e0135265. Available from: https://dx.plos.org/10.1371/journal.pone.0135265
45. NCI-60 DTP. NCI-60 Screening Methodology | NCI-60 Human Tumor Cell Lines Screen | Discovery & Development Services | Developmental Therapeutics Program (DTP) https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm [Internet]. [cited 2019 Dec 30]. Available from: https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm
46. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science (80-). 1975;187:226–32.
47. Riggs AD. X inactivation, differentiation, and DNA methylation. Cytogenet Genome Res. 1975;14:9–25.
48. Holliday R. A new theory of carcinogenesis. Br J Cancer. 1979;40:513–22.

49. Holliday R. Epigenetics: A Historical Overview. Epigenetics [Internet]. [cited 2020 Feb 1];1:76–80. Available from: http://www.landesbioscience.com/journals/epigenetics/abstract.php?id=2762
50. Cosgrove MS, Wolberger C. How does the histone code work? Biochem Cell Biol. 2005. page 468–76.
51. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature. 1983;301:89–92.
52. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell. 2011. page 646–74.
53. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell [Internet]. 1993;75:843–54. Available from: http://www.cell.com/article/009286749390529Y/pdf
54. Pal AS, Kasinski AL. Animal Models to Study MicroRNA Function. Adv Cancer Res. Academic Press Inc.; 2017. page 53–118.
55. Moutinho C, Esteller M. MicroRNAs and Epigenetics. Adv Cancer Res. Academic Press Inc.; 2017. page 189–220.
56. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37:391–400.

APPENDIX A. LCN2 OVEREXPRESSION SENSITIZES NON-SMALL CELL LUNG CANCER CELLS TO ERLOTINIB

Chapter overview

This chapter describes the process of identification of a novel protein, Lipocalin-2 (LCN2) that functions as a mediator of erlotinib response in non-small cell lung cancer (NSCLC). To identify the mechanistic role of LCN2 in mediating erlotinib response are further highlighted here.

A.1 Introduction:

A.1.1 Lipocalin 2 in cancer:

Lipocalin-2 (LCN2) is a 24kDa glycoprotein that belongs to the superfamily of Lipocalins. Lipocalins are small lipid transporters that are involved in the regulation of immune response and maintenance of cellular homoeostasis. In multiple studies, the role of LCN2 has been reported to be context dependent – based on the cell type, tumor type and the model system. For example, the loss of LCN2 has been reported to drive inflammation and tumorigenesis in an IL-10 knock-out model of colitis (1). Whereas, high levels of LCN2 is positively correlated with radio-resistance in oral squamous cell carcinoma (OSCC) and non-small cell lung cancer cells (NSCLC) cells (2), and increased invasiveness in breast cancer cells (3). In endometrial cancer, LCN2 has been implicated as an oncogene, inducing tumor progression by enhancing cell survival, preventing apoptosis, and increasing cell migration (4,5).

Contrastingly, LCN2 is also known to suppress certain hallmarks of cancer functioning as a tumor-suppressor. For instance, in OSCC, high LCN2 expression is correlated with better survival of patients (6), and loss of LCN2 in oral cancer cells is associated with increased survival, proliferation, migration and resistance to chemotherapeutics (7). Additionally, high levels of LCN2 suppresses invasion and angiogenesis in pancreatic cancer (8), and inhibits epithelial-to-mesenchymal transition (EMT) and metastasis in colorectal and hepatocellular cancer (9,10). LCN2 has also been shown to induce apoptosis in (NSCLC) (11,12). These functions of LCN2 suggest that it functions as a tumor suppressor in a context-dependent manner. Due to these controversial reports regarding LCN2 function in tumorigenesis and progression of cancers, and

its suggestive role in resistance, we aim to evaluate the role of LCN2 as a mediator of erlotinib response or resistance in NSCLC.

A.1.2 Targeted therapeutics for treatment of NSCLC:

Currently, upon diagnosis of NSCLC tumors, patients are treated with interventions that include surgery, chemotherapy, radiotherapy, targeted therapy or immunotherapy, depending on 1) the stage of the disease, and 2) overall health of the patient (13,14). Since most NSCLC patients are diagnosed with cancer in advanced stages, the routine is to profile the tumors to identify a driver mutation that is causal of the cancer (13,14). Identification of drivers subsequently aid in determination of the appropriate targeted therapies (**Figure 1.5, 1.6**, Chapter 1). In patients that are profiled for constitutively activated mutations in Epidermal Growth Factor Receptor (EGFR), are treated with the current standard of care, EGFR-Tyrosine Kinase Inhibitors (EGFR-TKIs). Although such patients display rapid response upon treatment with EGFR-TKIs, such as erlotinib, development of resistance that occurs in over 50% of cases and remains to be a major challenge. It is well reported that in over 60% of cases that develop erlotinib resistance, EGFR displays a secondary mutation, T790M and the remaining 20% tumors activate bypass tracks to circumvent their survival dependence on EGFR (**Figure 1.8**, Chapter 1). However, in 15-20% of cases, the mechanism of development of erlotinib resistance is largely unknown (**Figure 1.8**, Chapter 1), requiring additional research to determine altered levels of unidentified molecules that mediate erlotinib resistance.

A.1.3 Study design and hypothesis

LCN2 has been evaluated for its contribution in resistance to therapies in OSCC and NSCLC (2,7). Therefore, based on some convincing preliminary bioinformatic results (**Figure A.1**), this study was conducted with the hypothesis that LCN2 functions to promote erlotinib sensitivity in NSCLC cell lines.

A.2 Methods:

A.2.1 Cell culture:

All cell lines utilized in the study were obtained from American Type Culture Collection (ATCC), cultured under standard conditions and were confirmed to be free of mycoplasma. Stable cell lines constitutively expressing LCN2 generated during the study were grown in RPMI media supplemented with 10% FBS and 1% Penicillin/Streptomycin cocktail along with 20µg/ml G418 selection antibiotic. All the cells were cultured at 37°C in 5% CO_2.

A.2.2 Drug Preparation for in vitro studies:

Erlotinib (S7786, Selleck Chemicals) was dissolved in DMSO to prepare 0.4 M stock solutions, which were aliquoted and stored in -80°C. Working concentration of the drug was 200 µM prepared in complete medium and diluted to different concentrations for *in vitro* experiments.

A.2.3 Knockdown and overexpression experiments:

To knockdown LCN2, 25nM of silencing RNAs targeting LCN2 (siLCN2) (Silencer Select LCN2 siRNA, Life Tech, 4392420), and/or 25nM silencing RNA control (sicont) (Silencer Select Negative Control #2 siRNA, Life Technologies, 4390846) were transfected in $4X10^5$ cells plated in 6-well plates, using RNAiMAX, as per manufacturer's instructions. The plasmid containing the open reading frame of LCN2 (CMV-promoter driven LCN2 overexpression, p-cLCN2) was a gift from Dr. Shun-Fa Yang, Chung Shan Medical University. It was forward transfected in $4X10^5$ Calu-6 cells or was linearized and 2ug was forward transfected into $4X10^5$ EKVX or H358 cells using lipofectamine 2000 (11-668-019, Thermo Fisher Scientific), as per manufacturer's instructions to finally generate EKVX clone 1-7 and 2-4, and H358 clones 1-6, 2-2 and 2-5 by clonally selecting single cells in 50 or 100µg/ml G418, respectively.

A.2.4 RNA isolation and Quantitative real time PCR (qRT-PCR):

$4X10^5$ cells were grown in 6 well plates, and total RNA was isolated after time points as specified in figure legends, using the miRneasy Kit (217004, Qiagen) according to the manufacturer's instruction. DNase I digestion (79254, Qiagen) was used in each RNA preparation to remove genomic DNA. RNA integrity was evaluated on a 1.5% agarose gel, and total RNA quantified using a nanodrop. cDNA was then synthesized using 1µg of total RNA isolated from cells using MiScript Reverse Transcriptase kit (218161, Qiagen), as indicated by the manufacturer's protocol. miScript SYBR Green PCR Kit (218073, Qiagen) was utilized as indicated by the manufacturer's protocol, to quantify target gene mRNA expression normalized to the housekeeping, GAPDH via qRT-PCR. Primers used for *LCN2* QT00028098, Qiagen), *GAPDH* (loading control) (QT00079247, Qiagen).

A.2.5 Western Blot:

Cells ($4X10^5$) were grown in 6 well plates, and lysates were isolated at time points as specified in figure legends, using RIPA buffer [Sodium chloride (150 mM), Tris-HCl (pH 8.0, 50mM), N P-40 (1 %), Sodium deoxycholate (0.5 %), SDS (0.1 %), ddH2O (up to 100 mL)] containing 1X protease inhibitor cocktail (PIA32955, Thermo Fisher Scientific). Protein quantification was performed using Pierce BCA Protein Assay kit. Equal amounts of protein lysate were loaded and resolved on a 12% Polyacrylamide Gel, transferred onto a polyvinylidene difluoride (PVDF) membrane, blocked using LI-COR buffer for 1 hour at room temperature, and incubated overnight in primary antibody at 4 °C. The primary antibody was detected using 1:800 IR 800CW secondary antibody (Li-COR), blots were scanned, and data quantified by using the Odyssey LI-COR imaging system and software. Antibodies used: rabbit LCN2 (44058S, Cell Signaling), mouse β-ACTIN (3700, Cell Signaling).

A.2.6 Erlotinib dose response

The protocol followed to evaluate erlotinib dose response of various NSCLC cells and cells generated in this study was as per the NCI-60 Cell Five-Dose Screen (NCI-60, DTP (15)). Briefly, Sulforhodamine B colorimetric assay (SRB assay) was performed by exposing cells to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control)

containing media for 72 hours. Post data normalization, as described in figure legends, GI50 erlotinib was calculated from the respective dose curves NCI-60, DTP (15)).

A.2.7 Bioinformatic analyses:

Correlation analysis between adherent lung cancer cell lines and erlotinib response conducted by the publicly available database, Cancer Therapeutics Response Portal v2 (CtRP v2) was retrieved(16). Oncomine (17) was utilized to analyze published datasets from the studies conducted by Yauch *et al* (18) and Barretina *et al* (19). Coexpression analysis determines correlation of a gene with erlotinib sensitivity, and Cancer Outlier Profile Analysis (COPA) rank determines gene expression value (20,21). Coexpression close to 1 indicates high correlation of the gene with erlotinib resistance. Low COPA value indicates low gene expression. In Yauch study, cell lines exhibiting GI50 < 2.0 μmol/L were classified as sensitive, GI50 >8.0 μmol/L as resistant and GI50 2-8 μmol/L as intermediate erlotinib sensitivities (18). In Barretina study, 24 compounds and 500 cell lines were tested. Dose response curves were generated for each and an activity area between the response curve and the zero dose effect level was calculated based on which cells were classified as resistant (GI50 value higher than activity area), sensitive (GI50 value lower than activity area) and intermediate sensitivity (GI50 value within the activity area) (19).

A.2.8 Statistical analysis:

All data were analyzed using GraphPad Prism version 7 software (GraphPad Software) and are presented as mean values ± SEM. Student's t-test or one-way ANOVA were the statistical analyses performed, as specified in the figure legends. P-value of < 0.05 was considered significant.

A.3 Results

A.3.1 LCN2 positively correlates with erlotinib sensitivity in cancer cell lines, including lung cancer cell lines:

The publicly available database, Cancer Therapeutics Response Portal v2 (CtRP v2) was utilized to retrieve *LCN2* expression levels in adherent lung cancer cell lines that were exposed to erlotinib (16). A negative correlation between the half-maximal growth inhibitory concentration (GI50) of

erlotinib and *LCN2* (Correlation = -0.330) was determined (**Figure A.1A**). Simultaneously, the online available database, Oncomine (17) was used to evaluate *LCN2* transcript levels in cell lines that were categorized as erlotinib resistant, erlotinib intermediate sensitive, and erlotinib sensitive groups. The datasets that were used for this analysis came from published studies conducted by Yauch *et al* (18) and Barretina *et al* (19). Coexpression of data (Yauch study = 0.669, Barretina study = 0.629) indicate high correlation with erlotinib sensitivity. Low COPA values in Yauch study (COPA = 6.440) suggest low levels of *LCN2* was determined in the cell lines. Both results trended towards high *LCN2* in the erlotinib sensitive groups relative to the erlotinib resistant groups (**Figure A.1B, C**), suggesting a possible role for LCN2 in mediating erlotinib response.

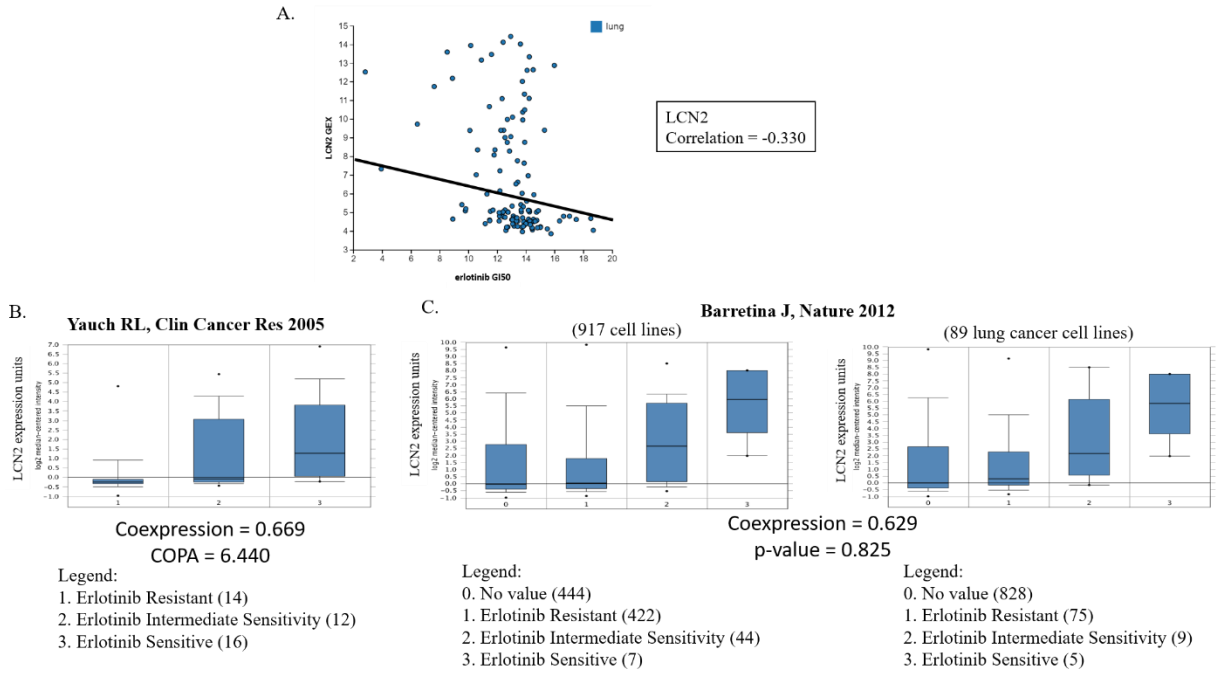


Figure A.1. LCN2 expression positively correlates with erlotinib sensitivity in multiple cell lines. A) Correlation analysis of *LCN2* expression levels in adherent lung cancer cell lines was retrieve from Cancer Therapeutics Response Portal v2 (CtRP v2). Erlotinib resistant, intermediate sensitivity and sensitive cell lines were retrieved using the Oncomine database and analysed relative to cell lines that were untreated. (A) Data was represented from Yauch RL (2005), and (B) Barretina (2012). "0. No values" in the legend are from cell lines that were not included in the study.

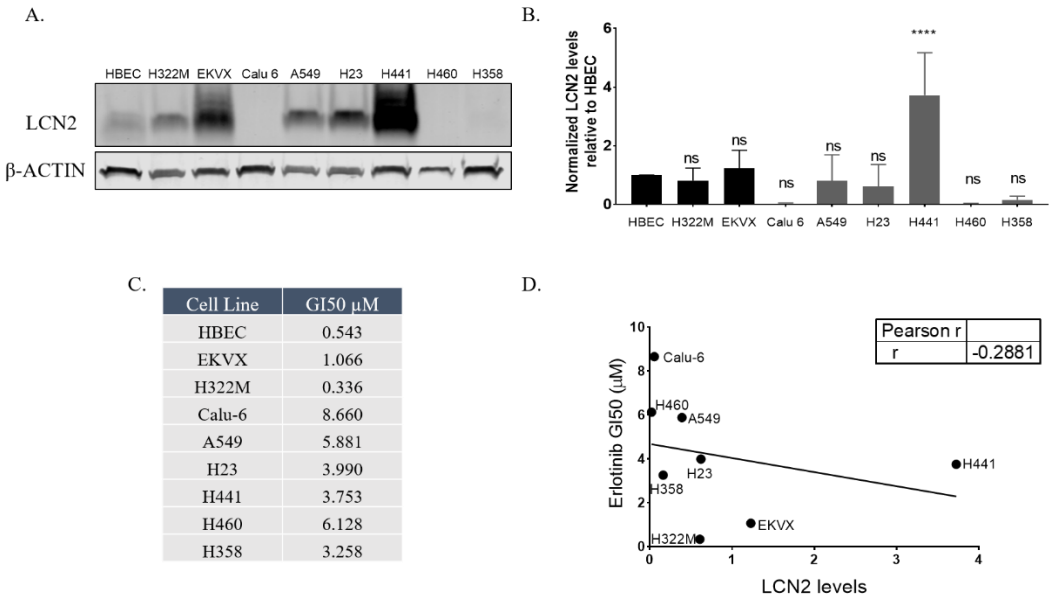
A.3.2 *LCN2* expression in NSCLC cells negatively correlates with erlotinib response

To evaluate if the endogenous levels of LCN2 in a panel of NSCLC cells in our laboratory also showed a similar correlation with erlotinib as evidenced from **Figure A.1A**, a western

blot analysis was performed (**Figure A.2A, B**) and were compared to the response of the cells to erlotinib (**Figure A.2C**). In the panel, both EKVX and H441 had the highest amount of cellular LCN2. While LCN2 was undetectable in Calu6, H460 and H358 cells. The 50% growth inhibitory concentration (GI50) of erlotinib for each of the NSCLC cell lines indicated that H322M and EKVX (not including Human Bronchial Epithelial Cells (HBEC) which are immortalized normal lung cells, and not cancer cell lines) were among the most sensitive, while Calu-6 and H358 were resistant (**Figure A.2C**). The correlation between LCN2 levels and the GI50 of erlotinib for each cell line was plotted and the results suggest that there is a negative correlation (Pearson's correlation = -0.2881) between LCN2 levels and GI50 erlotinib in NSCLC cells (**Figure A.2D)**.

A.3.3 Silencing LCN2 does not alter erlotinib response, but overexpressing LCN2 sensitizes NSCLC cells to erlotinib

Since LCN2 negatively correlates with erlotinib response, we next sought to investigate if modulation of LCN2 can alter erlotinib resistance. Two erlotinib sensitive cells, EKVX and H322M (48), were transfected with silencing RNAs targeting LCN2 and were monitored for their response to erlotinib . Although LCN2 was markedly downregulated (**Figure A.3A**), there was no apparent effect on erlotinib response in either of the NSLC cells (**Figure A.3B, C**). On the other hand, transiently overexpressing LCN2 in Calu-6 cells, sensitized the resistant Calu6 cells to erlotinib (**Figure A.3D, E**).



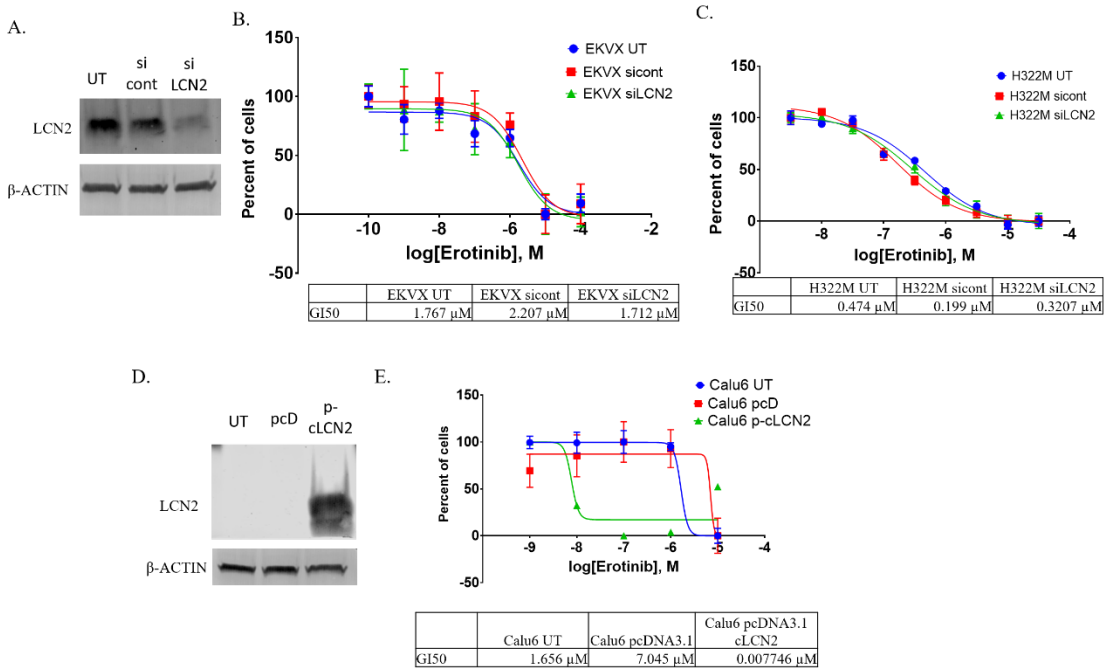
Cell Line	GI50 μM
HBEC	0.543
EKVX	1.066
H322M	0.336
Calu-6	8.660
A549	5.881
H23	3.990
H441	3.753
H460	6.128
H358	3.258

Figure A.2. LCN2 levels negatively correlates with erlotinib resistance in NSCLC cells. A) Representative western blot image of LCN2 in a panel of NSCLC cells. Human Bronchial Epithelial Cells (HBEC) is a normal lung epithelial cell line. β-ACTIN was utilized as a loading control. Western blot was performed using lysates isolated after 48 hours of culturing in complete media. B) Quantification of five biological replicates of western blot data. One-way ANOVA was used to evaluate the statistical significance of LCN2 levels relative to HBEC. C) Erlotinib dose response via SRB assay was evaluated by exposing each individual cell line in the panel of NSCLC cells to varying concentrations of erlotinib or the highest equivalent volume of DMSO (negative control) containing media for 72 hours. For percent of cells calculation, number of cells at time zero (tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. D) Correlation analysis between LCN2 levels from B and GI50 erlotinib concentrations from C for each cell line. A Pearson correlation test was conducted.

A.3.4 Stably overexpressing LCN2 in resistant NSCLC cells sensitizes them to erlotinib but induces cell death in sensitive cells

To further validate the effects of transient induction of LCN2 on NSCLC cells, especially re-sensitizing resistant cells, we generated single clones stably expressing the p-cLCN2 (constitutively overexpressing ORF of LCN2) plasmid, in two NSCLC lines and evaluated their erlotinib responses. The two cell lines utilized were: 1) EKVX, an erlotinib sensitive cell line, 2) H358, an intermediate erlotinib responsive cell line (not reported to be resistant or sensitive to erlotinib in literature). LCN2 overexpressing EKVX clones so generated post p-cLCN2 stable transfection were evaluated for LCN2 expression via qRT-PCR (**Figure A.4A**). Select EKVX

clones, clone 1-7 and 2-4 were further re-evaluated using western blot analysis (**Figure A.4A inset**). Dose curve analysis of these clones were not possible because there was massive cell death upon exposure to erlotinib (**Figure A.4C**). However, among H358 clones that were stably transfected with p-cLCN2 and clonally selected, evaluated via western blot (**Figure A.4B)**, clones 1-6, 2-2 and 2-5 were taken forward for dose curve analysis. Upon exposure to erlotinib, clones 2-2 and 2-5 showed marked sensitization to erlotinib, but clone 1-6 did not. It is possible that the plasmid constitutively expressing the ORF of LCN2 (p-cLCN2) was incorporated at a gene locus that altered erlotinib response (**Figure A.4D**).

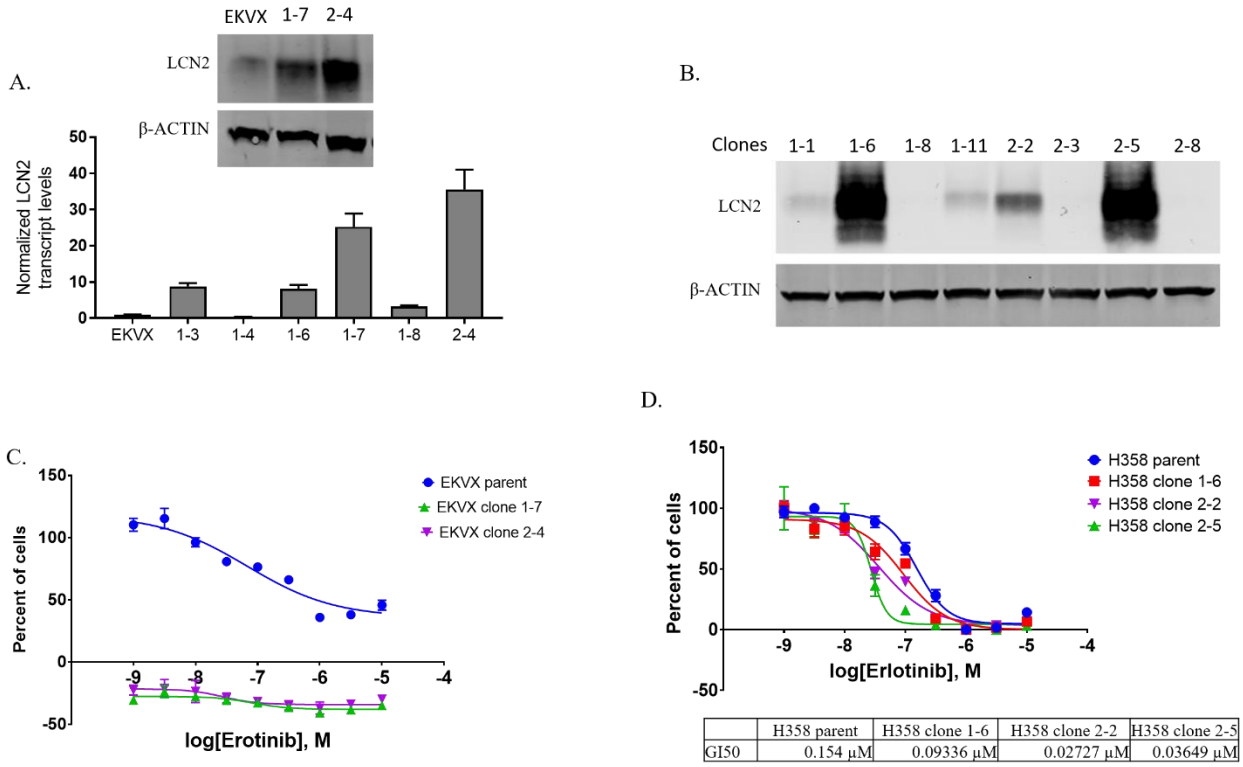


	EKVX UT	EKVX sicont	EKVX siLCN2
GI50	1.767 µM	2.207 µM	1.712 µM

	H322M UT	H322M sicont	H322M siLCN2
GI50	0.474 µM	0.199 µM	0.3207 µM

	Calu6 UT	Calu6 pcDNA3.1	Calu6 pcDNA3.1 cLCN2
GI50	1.656 µM	7.045 µM	0.007746 µM

Figure A.3. Overexpressing LCN2 sensitizes NSCLC cells to erlotinib. A) EKVX cells were either untransfected (UT), transfected with siRNA control (sicont), or transfected with siLCN2. Seventy-two hours post transfection, lysates were isolated and western blot performed for LCN2. β-ACTIN was utilized as a loading control. For transfected B) EKVX cells or C) H322M cells, erlotinib dose response was evaluated by exposing cells to varying concentrations of Erlotinib or the highest equivalent volume of DMSO (vehicle control) containing media for 72 hours. For percent of cells calculation, number of cells at the time of addition of Erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 erlotinib was calculated from the respective dose curves (as per NCI-60 Cell Five-Dose Screen, NCI-60, DTP (48)). D) Calu6 cells were either untransfected (UT), transfected with pcDNA3.1 control (pcD), or transfected with a p-cLCN2 (constitutively overexpressing ORF of LCN2). Seventy-two hours post transfection, lysates were isolated and western blot performed for LCN2. β-ACTIN was utilized as a loading control. E) For transfected Calu6 cells, erlotinib dose response was evaluated as described in B and C.



	H358 parent	H358 clone 1-6	H358 clone 2-2	H358 clone 2-5
GI50	0.154 µM	0.09336 µM	0.02727 µM	0.03649 µM

Figure A.4. Stably overexpressing LCN2 in NSCLC cells sensitizes them to erlotinib and induces cell death in sensitive cells. A) EKVX cells were transfected with cDNA of LCN2 expressing plasmid (p-cLCN2) to generate stably expressing LCN2 clones, evaluated for *LCN2* via qRT-PCR. LCN2 transcript levels are normalized to *GAPDH* and are graphed relative to expression in EKVX cells. Inset image depicts are representatie western blot image of clones 1-7 and 2-4. Lysates of EKVX (parental) or clones 1-7, 2-4 were isolated and western blot performed for LCN2. β-ACTIN was used as a loading control. B) Lysates of H358 clones generated by stably expressing p-cLCN2, were isolated and western blot performed for LCN2. β-ACTIN was utilized as a loading control. C) EKVX clones 1-7 and 2-4 and D) H358 clones 1-6, 2-2 and 2-5, where exposed to increasing doses of erlotinib for 72 hours and dose response curves were generated. For percent of cells calculation, number of cells at the time of addition of Erlotinib or DMSO (i.e. time zero or tz) was first corrected for, followed by normalization of cell number to respective corrected DMSO values. GI50 erlotinib was calculated from the respective dose curves (as per NCI-60 Cell Five-Dose Screen, NCI-60, DTP (48)).

A.4 Discussion and future directions:

Our findings corroborate with the results retrieved from two publicly available databases suggesting high levels of LCN2 correlate with erlotinib sensitivity (CtRP v2, Oncomine, **Figure A.1**). From the results in **Figures A.3** and **A.4**, it can be established that overexpressing LCN2 sensitizes cells to erlotinib. The findings also suggest that LCN2

levels in NSCLC is responsible for only sensitizing cells to the drug erlotinib, and even dramatically reducing LCN2 does not induce resistance in erlotinib sensitive cells (**Figure A.3, A.4**). Moreover, since LCN2 overexpression has been reported in multiple studies to induce cell death via apoptosis (11,22,23), it is not surprising that high levels of LCN2 induced cell death in erlotinib sensitive EKVX cells, when in combination with erlotinib (**Figure A.4A**). Therefore, it is required to be validated if stably overexpressing LCN2 truly induced apoptosis in these cells, by evaluating apoptosis markers. If apoptosis is initiated due to high levels of LCN2, it can be hypothesized that 1) LCN2 directly binds to apoptotic molecules and stabilizes/activates them, or 2) functions as a transcription factor to enhance transcription of the molecules involved. To identify if LCN2 directly initiates apoptotic molecules by physically interacting with them, an immunoprecipitation (IP) analysis post LCN2 overexpression and silencing can be conducted. Followed by evaluation of a few apoptosis mediators arrayed in the Human Apoptosis Antibody Array Kit. Comparing the results can provide a list of differentially pulled-down proteins that are possible interacting partners of LCN2, involved in mediating apoptotic response via LCN2. To evaluate if LCN2 functions as a potential transcription factor to transcriptionally activate molecules in the apoptotic pathway, an RNAseq can be performed in cells either overexpressing or knocked-down for LCN2. The differentially increased transcripts of the apoptotic pathway in LCN2 overexpression condition, can serve as candidates for further *in vitro* assessment.

Overall, the study identified a novel mechanism depicting that high levels of LCN2 can sensitize resistant NSCLC cells to erlotinib. Furthermore, high *LCN2* levels in multiple cell lines is also associated with sensitivity to other EGFR-TKIs other than erlotinib, such as afatinib and gefitinib (16). This lends stronger evidence to the findings of this study that indeed high levels of LCN2 may function as a mediator of erlotinib response. However, further *in vitro* and *in vivo* investigation of LCN2 in facilitating cell death in the presence of TKIs is warranted to determine that LCN2 is a *bona fide* mediator of erlotinib, and other TKI response in NSCLC. This can potentially translate into the use of LCN2 as a combinatorial therapeutic molecule with erlotinib, to sensitize erlotinib resistant tumors.

A.5 References:

1. Moschen AR, Gerner RR, Wang J, Klepsch V, Adolph TE, Reider SJ, *et al.* Lipocalin 2 Protects from Inflammation and Tumorigenesis Associated with Gut Microbiota Alterations. Cell Host Microbe **2016**;19:455-69
2. Shiiba M, Saito K, Fushimi K, Ishigami T, Shinozuica K, Nakashima D, *et al.* Lipocalin-2 is associated with radioresistance in oral cancer and lung cancer cells. International Journal of Oncology **2013**;42:1197-204
3. Fougere M, Gaudineau B, Barbier J, Guaddachi F, Feugeas JP, Auboeuf D, *et al.* NFAT3 transcription factor inhibits breast cancer cell motility by targeting the Lipocalin 2 gene. Oncogene **2010**;29:2292-301
4. Lin HH, Liao CJ, Lee YC, Hu KH, Meng HW, Chu ST. Lipocalin-2-induced cytokine production enhances endometrial carcinoma cell survival and migration. Int J Biol Sci **2011**;7:74-86
5. Mannelqvist M, Stefansson IM, Wik E, Kusonmano K, Raeder MB, Oyan AM, *et al.* Lipocalin 2 expression is associated with aggressive features of endometrial cancer. BMC Cancer. England2012. p 169.
6. Lin CW, Yang WE, Lee WJ, Hua KT, Hsieh FK, Hsiao M, *et al.* Lipocalin 2 prevents oral cancer metastasis through carbonic anhydrase IX inhibition and is associated with favourable prognosis. Carcinogenesis **2016**;37:712-22
7. Monisha J, Roy NK, Padmavathi G, Banik K, Bordoloi D, Khwairakpam AD, *et al.* NGAL is Downregulated in Oral Squamous Cell Carcinoma and Leads to Increased Survival, Proliferation, Migration and Chemoresistance. Cancers **2018**;10:228
8. Tong Z, Kunnumakkara AB, Wang H, Matsuo Y, Diagaradjane P, Harikumar KB, *et al.* Neutrophil gelatinase-associated lipocalin: a novel suppressor of invasion and angiogenesis in pancreatic cancer. Cancer Res **2008**;68:6100-8
9. Feng M, Feng J, Chen W, Wang W, Wu X, Zhang J, *et al.* Lipocalin2 suppresses metastasis of colorectal cancer by attenuating NF-kappaB-dependent activation of snail and epithelial mesenchymal transition. Mol Cancer **2016**;15:77

10. Wang YP, Yu GR, Lee MJ, Lee SY, Chu IS, Leem SH, *et al.* Lipocalin-2 negatively modulates the epithelial-to-mesenchymal transition in hepatocellular carcinoma through the epidermal growth factor (TGF-beta1)/Lcn2/Twist1 pathway. Hepatology **2013**;58:1349-61
11. Hsin IL, Hsiao Y-C, Fang M-F, Jan M-S, Tang S-C, Lin Y-W, *et al.* Lipocalin 2, a new GADD153 target gene, as an apoptosis inducer of endoplasmic reticulum stress in lung cancer cells. Toxicology and Applied Pharmacology **2012**;263:330-7
12. Wu M-F, Hsiao Y-M, Huang C-F, Huang Y-H, Yang W-J, Chan H-W, *et al.* Genetic Determinants of Pemetrexed Responsiveness and Nonresponsiveness in Non-small Cell Lung Cancer Cells. Journal of Thoracic Oncology **2010**;5:1143-51
13. Lung Cancer 101. **2014**
14. 2020 Lung Cancer 101 | Lungcancer.org. <https://www.lungcancer.org/find_information/publications/163-lung_cancer_101/269-non_small_cell_lung_cancer_treatment>.
15. NCI-60 D. 2020 NCI-60 Human Tumor Cell Lines Screen | Discovery & Development Services | Developmental Therapeutics Program (DTP). <https://dtp.cancer.gov/discovery_development/nci-60/>.
16. 2020 Cancer Therapeutics Response Portal. <https://portals.broadinstitute.org/ctrp.v2.1/?compoundId=52928&compoundName=undefined>.
17. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, *et al.* ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia **2004**;6:1-6
18. Yauch RL, Januario T, Eberhard DA, Cavet G, Zhu W, Fu L, *et al.* Epithelial versus Mesenchymal Phenotype Determines In vitro Sensitivity and Predicts Clinical Activity of Erlotinib in Lung Cancer Patients. **2005**
19. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature **2012**;483:603-7
20. Tomlins SA, Rhodes DR, Yu J, Varambally S, Mehra R, Perner S, *et al.* The role of SPINK1 in ETS rearrangement-negative prostate cancers. Cancer Cell **2008**;13:519-28

21. Oncomine. 2020 <https://cbt.bio.uci.edu/files/2016/01/Oncomine-Manual_Module2_v43.pdf>.
22. Tang W, Department of Ophthalmology EaEHoFU, Shanghai, China, Ma J, Research Center EaEHoFU, Shanghai, China, Gu R, Department of Ophthalmology EaEHoFU, Shanghai, China, *et al.* Light-Induced Lipocalin 2 Facilitates Cellular Apoptosis by Positively Regulating Reactive Oxygen Species/Bim Signaling in Retinal Degeneration. Investigative Ophthalmology & Visual Science **2020**;59:6014-25
23. Xu G, Ahn J, Chang S, Eguchi M, Ogier A, Han S, *et al.* Lipocalin-2 induces cardiomyocyte apoptosis by increasing intracellular iron accumulation. J Biol Chem **2012**;287:4808-17

VITA

Arpita Sushant Pal

March 2020

Graduate Research Assistant, Dr. A. Kasinski Laboratory;
Department of Biological Sciences, Bindley Multidisciplinary Cancer Research Center, Room 281,
West Lafayette, Indiana 47906

EDUCATION

Ph.D. [Present] **Purdue University**. West Lafayette, IN. Area: Interdisciplinary Life Sciences.
Focus: Cell and Molecular Biology Cluster
Advisor: Dr. Andrea Kasinski

M.Sc. [2011] **University of Sheffield**. Sheffield, UK. Area: Molecular Medicine.
Focus: Experimental Medicine
Thesis: Carcinogen-induced modulation of microRNA expression profiles in oral epithelium.
Advisor: Dr. Daniel Lambert

B.Sc. [2010] **Sinhgad Technical Education Society's College of Science**, Pune, India.
Area: Biotechnology
Thesis: Comparison of antioxidant levels, genomic and mitochondrial DNAs of diploid and tetraploid *Spinacia oleracea* L.
Advisor: Dr. Suman Sheelavantmath

PROFESSIONAL ACTIVITIES

Fall 2019	Bilsland Dissertation Fellowship, Ph.D. Candidate, Department of Biological Science, Purdue University (Dr. Andrea Kasinski)
2018 - 2019	Cancer Prevention Internship Program (CPIP) Graduate Research Assistantship, Purdue University
2017 - 2018	SIRG Graduate Research Assistant, Purdue University
2016 - 2017	Purdue Research Foundation (PRF) Research Assistant
2014 - 2015	Graduate Research Assistant, Purdue University (Dr. Andrea Kasinski)
2013 - 2014	Graduate Research Assistant, Purdue University (Purdue University Life Sciences Program (PULSe))

2012 - 2013 Voluntary Research Assistant, Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Mumbai, India

2011 - 2012 Visiting Researcher, University of Sheffield, Sheffield, UK

AWARDS AND HONORS

2019 Bilsland Dissertation Fellowship (Fall 2019), Department of Biology, Purdue University.

2018 The Center for Cancer Research Miles Graduate Scholarship Award ($1200),
Purdue University

2018 Cancer Prevention Internship Program Graduate Research Assistantship, Purdue
University.

2018 Executive Vice President for Research and Partnership (EVPRP) grant to obtain the chemotaxis Software Module and Cell Migration Kit for Incucyte S3 ($21,350).

2017 SIRG Graduate Research Assistantship, The Purdue University Center for Cancer Research, Purdue University.

2017 Featured in Indiana University News article – *http://www.cancer.iu.edu/news-publications/cancer-research-day.shtml*

2017 ***Awarded second place***- Aberrantly expressed microRNAs drive the development of acquired Erlotinib resistance in non-small cell lung cancer. Poster competition. The IU Simon Cancer Center's Cancer Research Day, Indiana University, Indianapolis.

2017 Yeunkyung Woo Achieve Excellence Travel Award 2017, Department of Biology, Purdue University.

2017 Purdue Graduate Student Government and the Graduate School Travel Grant, Purdue University.

2017 Purdue University Center for Cancer Research Travel Award 2017, Purdue University.

2016 PULSe Travel Award 2017, Purdue University.

2016 Purdue Research Foundation (PRF) Research Grant award, Department of Biology, Purdue University.

2016 ***Awarded first place*** - MicroRNAs: altering response to chemotherapy in lung cancer. 5-minute Thesis Competition, Purdue University Interdisciplinary Life Sciences Welcome Reception, Purdue University. West Lafayette, IN.

2013 Best poster award- Carcinogen-induced modulation of microRNA expression profiles in oral epithelium. Indian Cancer Genetics Conference. ACTREC, Mumbai, India.

2007-2010 Consistent top rank holder in B.Sc. Biotechnology, Sinhgad Technical Education Society's College of Science, Pune, India.

TEACHING AND MENTORSHIP

2017 - 2019	Mentored one undergraduate student (Fall 2017 to Fall 2019).
2017	Guest Lecturer for microRNAs, small RNAs with a big role, Epigenetics and Human Disease (BIOL 59500), Purdue University.
2017	Mentored one high school student (Summer 2017).
2016 - 2017	Mentored one undergraduate student (Fall 2016 to Spring 2017).
2016	Mentored one exchange student from The National University of Columbia, Bogota, Columbia (Spring 2016).
2016	Laboratory in Biology IV Genetics and Molecular Biology, Spring 2016, Teaching Assistant, Purdue University.
2016	Mentored one high school student (Summer 2016).
2012	Primary cells and cell lines, Animal Tissue Culture, Visiting Lecturer, Sinhgad Technical Education Society's College of Science, Pune, India.

SERVICE AND AFFILIATIONS

2019	Volunteer at Nanodays, Purdue University
2019	Volunteer at Relay for Life (CPIP), Purdue University
2019	Volunteer at Spring Fest (CPIP), Purdue University
2019	Host and volunteer at skin cancer prevention movie night and panel discussion (CPIP), Purdue University
2019	Host and volunteer at skin cancer awareness booth – pamphlet and sunglasses distribution (CPIP), Purdue University
2017 - Present	Science Associate at STEM innovations (http://www.steminnovations.org/)
2016 - Present	Member at American Association for Cancer Research
2016 - Present	Member at the RNA Society
2014 - Present	Member at Association for Women in Science
2014	Volunteer at Science in Schools, Purdue University

PUBLICATION INFORMATION

Peer Reviewed Publications:

Pal, A.S., Agredo, A., Lanman, N. A., Clingerman, J., Gates, K., Kasinski, A. L. (2020). Loss of SUV420H2 promotes EGFR inhibitor resistance in NSCLC through upregulation of MET via LINC01510. bioRxiv.03.17.995951; doi: https://doi.org/10.1101/2020.03.17.995951

Pal, A.S., Agredo, A., Kasinski, A.L. (2020). In-cell western protocol for semi high throughput screening of single clones. BioProtocols (accepted)

Zhou, W., **Pal, A.S**., Hsu, A. Y-H., Gurol, T., Zhu, X., Wirbisky, S.E., Freeman, J.L., Kasinski, A.L., Deng, Q. (2018). MicroRNA-223 suppresses the canonical NF-κB pathway in basal keratinocytes to dampen neutrophilic inflammation. Cell Reports, 13;22(7):1810-1823.

Pal, A., Melling, G., Hinsley, E.E., Kabir, T.D., Colley, H.E., Murdoch, C., Lambert, D.W. (2012). Cigarette smoke condensate promotes pro-tumourigenic stromal-epithelial interactions by suppressing miR-145. Journal of Oral Pathology & Medicine, 42(4):309-14.

Review Article:
Pal, A. S., & Kasinski, A. L. (2017). Animal Models to Study MicroRNA Function. Advances in Cancer Research, 135, 53-118.

Conference Abstract:
Arpita S. Pal, Alejandra Agredo, Andrea L. Kasinski. Aberrantly expressed microRNAs drive the development of acquired Erlotinib resistance in non-small cell lung cancer [abstract]. In: Proceedings of the American Association for Cancer Research Annual Meeting 2017; 2017 Apr 1-5; Washington, DC. Philadelphia (PA): AACR; Cancer Res 2017;77(13 Suppl): Abstract nr 3142. doi:10.1158/1538-7445.AM2017-3142

PRESENTATIONS
Pal, A. S., Agredo, A., Bains, M., Kasinski, A. 2019. Epigenetic mediators of Erlotinib resistance in Non-Small Cell Lung Cancer. 3rd International Conference on the Long and the Short of Non-Coding RNAs, Aegean Conferences, Crete, Greece.
Pal, A. S., Agredo, A, Kasinski, A. 2018. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. The 2018 Midwest Chromatin and Epigenetics Meeting, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Biology Retreat, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Purdue University Interdisciplinary Life Sciences Program (PULSe) Retreat, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. IU Simon Cancer Center's Cancer Research Day. Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN ***(Awarded second place)***
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Purdue's OIGP Spring Reception, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Postdoc/Senior Graduate Student Symposium, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. AACR 2017. Washington, D.C.
Pal, A. S., Agredo, A, Kasinski, A. 2016. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Chromatin and

Epigenetics Symposium, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2016. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. 3rd Annual Drug Discovery Symposium, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2016. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Purdue Cancer Research Day, Purdue University.
Pal, A. S., Kasinski, A. 2015. Involvement of Micrornas In Regulation of Lipocalin2 In Non-Small Cell Lung Carcinomas. Purdue Cancer Research Day, Purdue University.
Pal, A. S., Kasinski, A. 2015. Involvement of Micrornas In Regulation of Lipocalin2 In Non-Small Cell Lung Carcinomas. Biology Graduate Student Recruitment Event, Purdue University.
Pal, A. S., Kasinski, A. 2015. Involvement of Micrornas In Regulation of Lipocalin2 In Non-Small Cell Lung Carcinomas. Purdue's OIGP Spring Reception, Purdue University.
Pal, A. S., Kasinski, A. 2014. Interplay Between LCN2, Mir-155 And Mir-21 In Non-Small Cell Lung Carcinomas. Purdue Center for Cancer Research Annual Scientific Retreat, Purdue University.
Pal, A. S., Andrasani, O. 2014. Ovarian cancer heterogeneity studied in 3-Dimension by Motility Contrast Imaging. Purdue University Interdisciplinary Life Sciences Program (PULSe) first-year poster competition, Purdue University.
Pal, A. S., Andrasani, O. 2014. Ovarian cancer heterogeneity studied in 3-Dimension by Motility Contrast Imaging. Purdue's OIGP Spring Reception, Purdue University.
Pal, A. S., Andrasani, O. 2013. Ovarian cancer heterogeneity studied in 3-Dimension by Motility Contrast Imaging, Phi Zeta Purdue University School of Veterinary Medicine Poster Competition, Purdue University.
Pal, A. S., Lambert, D. 2013. Carcinogen-induced modulation of microRNA expression profiles in oral epithelium. Indian Cancer Genetics Conference. ACTREC, Mumbai, India***(Awarded first place)***

ORAL PRESENTATIONS
Pal, A. S., Agredo, A., Bains, M., Kasinski, A. 2019. Epigenetic mediators of Erlotinib resistance in Non-Small Cell Lung Cancer. Cell and Molecular Biology Cluster, Department of Biological Sceinces, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2018. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. The Hitchhiker's Guide to the Biomolecular Galaxy, Purdue University.
Pal, A. S., Agredo, A, Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. 2- minute talk at Biology Retreat, Purdue University.
Pal, A. S., Agredo, A., Kasinski, A. 2017. Aberrantly expressed microRNAs drive the development of acquired Erlotinib-resistance in Non-Small Cell Lung Cancer. Cell and Molecular Biology Cluster, Department of Biological Sceinces, Purdue University.

Pal, A. S., Kasinski, A. 2016. MicroRNAs: altering response to chemotherapy in lung cancer. 5- minute Thesis Competition, Purdue University Interdisciplinary Life Sciences Welcome Reception, Purdue University. West Lafayette, IN ***(Awarded first place)***

Pal, A. S., Sheelavantmath, S. 2010. Comparison of antioxidant levels, genomic and mitochondrial DNAs of diploid and tetraploid *Spinacia oleracea* L. 3rd State level seminar on Agricultural Biotechnology for sustainable productivity, Amravati, India.

SKILLS

Skilled at performing sterile cell culture techniques, cell-based assays – proliferation assays, scratch assays, migration assays, luciferase assays, molecular biology techniques such as western blots, in-cell westerns, immunofluorescence, PCRs, molecular cloning, lentiviral library generation.

Skilled at conducting *in vivo* experiments that include mouse injections, dissections and imaging.

Proficient with bioinformatic softwares such as Galaxy, IPA, TCGA, MAGeCK-VISPR, UCSC Genome Browser, NCBI, DAVID, DIANA-miRPathv3.0, GSEA, Prism, JMP graphing software.

Experienced in basic use of computer languages such as Linux, R (course: Intro to R and Bioconductor (BCHM 69500)) and Python (Coursera online course: Python for Everybody).