

UNIVERSITY OF CALIFORNIA SAN DIEGO

GRAS1 Long Non-coding RNA Protects Cells from Ferroptosis and DNA Damage

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Biochemistry and Molecular Biophysics

by

Tong Su

Committee in charge:

Professor Colleen McHugh, Chair
Professor Daniel Donoghue
Professor Mark Young
Professor Jing Yang
Professor Brian Zid

2023

Copyright

Tong Su, 2023

All rights reserved.

The Dissertation of Tong Su is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

This work is dedicated to my dearest parents who have always supported me for every aspect of my life. I would not be where I am today without their support. Thank you so much!

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE	iii
DEDICATION	iv
TABLE OF CONTENTS.....	v
LIST OF FIGURES	vii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS.....	x
ACKNOWLEDGEMENTS	xi
VITA.....	xiv
ABSTRACT OF THE DISSERTATION	xv
CHAPTER1: Introduction.....	1
1.1 Definition and Classification of lncRNA	1
1.2 lncRNA as biomarkers for disease diagnosis and prognosis.....	2
1.3 Role of lncRNA in cancer	3
1.4 lncRNA databases and resources.....	5
1.5 lncRNA GRAS1	6
1.6 Lung cancer	7
1.7 DNA damage and p53 signaling pathway	8
1.8 RNA-antisense purification with mass spectrometry	9
1.9 Reference	11
CHAPTER2: Exploring functions of novel cancer-associated long non-coding RNAs in cancer cells	18
2.1 Abstract.....	18
2.2 Introduction	19
2.3 Result.....	23
2.4 Discussion/future directions	37
2.5 Reference	39
2.6 Acknowledgement.....	43
CHAPTER3: GRAS1 long non-coding RNA directly interacts with STAT5 to regulate STAT5 activity	44
3.1 Abstract.....	44
3.2 Introduction	44

3.3 Result.....	47
3.4 Discussion/future directions	58
3.5 Reference	60
3.6 Acknowledgement.....	65
CHAPTER4: GRAS11, a DNA-inducing lncRNAs, regulates cell survival by interacting with NKAP in lung cancer cell	66
4.1 Abstract.....	66
4.2 Introduction	67
4.3 Result.....	71
4.4 Discussion/future directions	99
4.5 Reference	102
4.6 Acknowledgement.....	108
CHAPTER 5: Materials and methods	109
5.1 Cell lines and culture conditions	109
5.2 Plasmid generation and transfection.....	109
5.3 Gene knockdown	110
5.4 RNA, DNA isolation and qRT-PCR	110
5.5 Immunoblot	111
5.6 Cellular fractionation.....	112
5.7 Colony formation assay	113
5.8 Immunofluorescence	113
5.9 Immunoprecipitation	114
5.10 MTT cell survival assay	115
5.11 Cell cycle analysis	116
5.12 RNA sequencing.....	116
5.13 RNAseq analysis	116
5.14 RNA antisense purification with mass spectrometry	117
5.15 Annexin V apoptotic assay	118
5.16 Ferroptosis assay	119

LIST OF FIGURES

Figure 2.1: Efficient overexpression of lncRNAs in K562 using pcDNA3.2 vectors. ..	28
Figure 2.2: Inefficient knockdown of lncRNAs in K562 using siRNA.	29
Figure 2.3: Efficient knockdown of lncRNAs in K562 using GapmeRs	30
Figure 2.4: Cell survival assay on the lncRNA targets	31
Figure 2.5: Altered expression of linc01977 affects multiple transcriptomic programs in MCF7 cells	33
Figure 2.6: Altered expression of GRAS1 affects multiple transcriptomic programs in A549 cells	34
Figure 2.7: Overview of probe design and RNA target capture efficiency with different probe length	35
Figure 2.8: Overview of workflow on testing probe capture and lncRNA target capture efficiency	36
Figure 3.1: Overview of workflow on RAP-qPCR and testing capture efficiency of the DNA probes	52
Figure 3.2: U1 binds to similar protein interacting partners in different cell lines	53
Figure 3.3: GRAS1 directly interacts with STAT5 in K562 cells.....	54
Figure 3.4: GRAS1 binds to STAT5 in K562 cells.....	55
Figure 3.5: Knockdown of GRAS1 affects phosphorylation of STAT5	56
Figure 3.6: Knockdown of GRAS1 affects STAT5 signaling pathway	56
Figure 3.7: Knockdown of GRAS1 affects subcellular localization of STAT5.....	57
Figure 4.1: GRAS1 is a well-expressed lncRNA located in both nucleus and cytoplasm in A549	83

Figure 4.2: GRAS1 can be efficiently overexpressed and downregulated by pcDNA vector and GapmeR respectively	84
Figure 4.3: GRAS1 is required for lung cancer cell survival	85
Figure 4.4: GRAS1 affects multiple transcriptomic programs in A549.....	86
Figure 4.5: Knockdown of GRAS1 activates p53 signaling pathway	87
Figure 4.6: Overexpression of GRAS1 induces DNA damage and activates p53 signaling pathway.....	88
Figure 4.7: GRAS1 affects cell cycle distribution.....	89
Figure 4.8: Knockdown of GRAS1 induces DNA damage.....	90
Figure 4.9: GRAS1 directly interacts with NKAP	91
Figure 4.10: NKAP binds with multiple ferroptosis-related RNAs	92
Figure 4.11: Downregulation of GRAS1 decreases NKAP level, which leads to a decrease in NF- κ B signaling pathway	93
Figure 4.12: GRAS1 post-translationally protects NKAP from ubiquitination	94
Figure 4.13: GRAS1 affects cellular ferroptosis	95
Figure 4.14: Overexpression of NKAP increase cell viability and protects GRAS1-lacking cells from ferroptosis	96
Figure 4.15: Overexpression of NKAP protects GRAS1-lacking A549 from DNA damage	97
Figure 4.16: Overexpression of NKAP rescues cell death caused by GRAS1 knockdown in A549 cells	98

LIST OF TABLES

Table 5.1: List of antibodies	119
Table 5.2: List of qPCR primers.....	120
Table 5.3: List of cell lines	120

LIST OF ABBREVIATIONS

lncRNA	Long Non-Coding RNA
mRNA	Messenger RNA
cDNA	Complementary DNA
RT	Reverse Transcription
eGFP	Enhanced Green Fluorescent Protein
WT	Wild Type
OE	Overexpression
KD	Knockdown
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
PCA	Principal Component Analysis
RAP	RNA Antisense Purification
MS	Mass Spectrometry
GRAS1	Growth Regulator Antisense 1
STAT5	Signal transducer and activator of transcription 5
NKAP	NF- κ B Activating Protein
IP	Immunoprecipitation
RIP	RNA Immunoprecipitation
RNAseq	RNA Sequencing
GO	Gene Ontology
SD	Standard Deviation

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my PhD research advisor, Dr. Colleen McHugh, for her support and guidance during this journey at UCSD. Exploring the field of long non-coding RNA, which people usually question a lot, is really challenging, yet exciting and rewarding at the same time. Dr. Colleen McHugh is always patient and encouraging to support my research, and I really appreciate that. During my PhD program in the McHugh Lab, I have learnt a lot not only the experimental ideas and techniques on RNA research, but also how to become a better scientist with self-learning skills and critical thinking.

I also want to thank my committee members (Dr. Brian Zid, Dr. Daniel Donoghue, Dr. Jing Yang and Dr. Mark Young) for giving me their valuable advice and guidance during my PhD program at UCSD. Thank you so much for your time and effort serving on my committee.

Certainly, I would like to appreciate and acknowledge my great McHugh lab mates. Firstly, I would like to appreciate my great hardworking students (Calvin Huang, Bobby Kong, Darin Cheung, and Jonathan Zhu) for helping me a lot with research. I definitely cannot achieve that much on my research projects without them. In addition, Aileen, Adarsh, Nhu, and Simone were all awesome friends in the lab, and they provided me with a lot of useful ideas and support for my research projects. Thank you for giving me such a wonderful time in the McHugh lab.

In addition, I would like to thank my undergraduate research advisor, Dr. Mark Hannink, for his guidance and encouragement to become a scientist. I still remember how

difficult it was to pursue a degree in science at Missouri as a new immigrant who is a non-English native speaker. However, Dr. Hannink deeply understood my difficulties and kept encouraging me not to give up on doing science. He invited me to join his lab and taught me a lot of experimental techniques hand by hand. In addition, he provided me with a lot of help in getting a scholarship, summer internship, and graduate school. I sincerely appreciated his support and guidance when I was an undergraduate student there. Thank you so much Dr. Hannink!

Finally, I would like to thank my parents for their endless support and love. Although neither of my parents have graduated from high school, they always encourage me to pursue science. Moving to the United States to pursue a degree in science was a big challenge for me. When I reached the US, I had to immerse myself and adapt to a culture that was different from what I was used to. The lack of friends, fluent English skills, and the coursework made me feel overwhelmed in the new environment. They completely understood my difficulties, and never stopped getting me the mental and financial support that I needed. Without them, it would be impossible for me to come here and pursue a PhD in the US. I deeply appreciate all the effort and sacrifice that they gave. Thank you, dad, and mom!

Chapter 2, in full, is currently being prepared for publication of the material. The authors in the publication: Tong Su, Calvin Huang, Lingbo Kong, Jonathan Zhu, and Colleen McHugh. The dissertation author is the primary investigator and author of this material.

Chapter 3, in full, is currently being prepared for publication of the material. The authors in the publication: Tong Su, Lingbo Kong, Darin Cheung, Calvin Huang, and Colleen McHugh. The dissertation author is the primary investigator and author of this material.

Chapter 4, in full, is currently being submitted for publication. The authors in the publication: Tong Su, Lingbo Kong, Darin Cheung, Jonathan Zhu, Nichelle Camden, Wei-Han Chou, Calvin Huang, and Colleen McHugh. The dissertation author is the primary author of this material. We thank Brian Zid (University of California San Diego) for assistance with confocal microscopy and flow cytometry. This work was supported by Startup Funding from university of California San Diego, NCI Cancer Training grant T32 CA009523 (Tong Su), University of California Cancer Research Coordinating Committee grant support C21CR2119, Undergraduate Summer Research Award (Lingbo Kong, Calvin Huang).

VITA

2017.12 Bachelor of Science in Biochemistry, University of Missouri
Columbia

2020.03 Master of Science in Chemistry and Biochemistry, University of
California San Diego

2023.06 Doctor of Philosophy in Biochemistry and Molecular Biophysics,
University of California San Diego

ABSTRACT OF THE DISSERTATION

GRAS1 Long Non-coding RNA Protects Cells from Ferroptosis and DNA Damage

by

Tong Su

Doctor of Philosophy in Biochemistry and Molecular Biophysics

University of California San Diego, 2023

Professor Colleen McHugh, Chair

Long non-coding RNAs (lncRNAs) have been implicated as regulators of cell growth and survival in a variety of cancer types, but the specific biological functions and interaction partners of many lncRNAs are still not well understood. In this dissertation, I used multiple biochemical and quantitative techniques to study how a novel cancer-associated lncRNA, termed Growth Regulator Antisense 1 (GRAS1), mechanistically regulates cell viability in cancer cells.

In Chapter 2, I discussed the development of a system to investigate 5 new cancer-associated lncRNAs from screening data from multiple published studies, including linc01977, linc00883, PDCD4-AS1, linc00909, and GRAS1. More specifically, we established a working platform of overexpression and knockdown for these lncRNAs, performed cell survival screening on these lncRNAs using five different cancer cell lines, examined transcriptomic alterations in altered expression of these lncRNAs, and validated biotinylated probes to capture these lncRNAs. Based on the data in that chapter, we selected GRAS1 to be our lncRNA of interest to focus on. In Chapter 3, I reported that GRAS1 directly and specifically interacts with STAT5 in K562 cells using RNA antisense purification with mass spectrometry (RAP-MS). In addition, we found that knockdown of GRAS1 reduces phosphorylation of STAT5 as well as the activity of the STAT5 signaling pathway in K562 cells. In Chapter 4, I characterized the function of GRAS1 in non-small cell lung cancer cells. GRAS1 is critical for A549 cell survival, and depletion of GRAS1 RNA transcript induces cell death. We found that knockdown of GRAS1 resulted in a significant increase in protein levels of p53 and the DNA damage marker γ H2AX. GRAS1 knockdown leads to differential expression of many mRNA transcript levels, including upregulation of DNA damage response factors involved in double strand break repair. Further, we identified the NF- κ B activating

protein (NKAP) as a direct interaction partner of GRAS1 using RAP-MS. Decreased GRAS1 levels in lung cancer cells caused a significant decline in NKAP level and induction of cell death through ferroptosis. The addition of the ferroptosis inhibitors ferrostatin and α -tocopherol could partially rescue the effect of GRAS1 knockdown on lung cancer cell survival. We found NKAP overexpression could completely reverse the ferroptosis as well as rescue the cell viability promoted by GRAS1 downregulation in lung cancer cells. Besides, NKAP overexpression could significantly reverse the effect of GRAS1 knockdown on lung cancer DNA damage. Chapter 4, in full, is currently being submitted for publication. Finally in Chapter 5, I listed the methods and materials we used in this study.

Chapter 1: Introduction

1.1 Definition and Classification of lncRNA

Recent high throughput sequencing studies have revealed that over 60% of the human genome is transcribed into RNA, representing a vast unexplored landscape of potential cellular regulatory molecules (Carninci et al. 2005). The number of expressed non-coding RNAs is correlated with increasing organismal complexity, while the number of protein-coding genes has remained relatively constant at around 20,000 in flies, mice, and humans over evolutionary time (Liu et al. 2013).

Long non-coding RNA, which is referred to as lncRNA, is a class of RNA molecules that are longer than 200 nucleotides without the capability of being translated into proteins (Wilusz et al. 2009). At the beginning, lncRNA was considered to be transcriptional noise, but numbers of recent studies have shown that they play critical roles in various biological processes, including cellular metabolism (Wang et al. 2017), cell proliferation (Zhang et al. 2018), cell differentiation (Fatica and Bozzoni 2014), and cell migration (Lingadahalli et al. 2018). LncRNAs are now known to regulate gene expression through diverse mechanisms, such as chromatin modification, transcriptional regulation, RNA processing, and post-transcriptional modification (Guttman and Rinn 2012).

LncRNAs are classified into various subtypes based on their genomic location, structure, and function. The genomic location-based classification includes intergenic, intronic, and antisense lncRNAs, which are transcribed from intergenic regions, introns, and the opposite strand of protein-coding genes, respectively (Shi et al. 2021). Structural classification includes linear, circular, and enhancer RNAs based on their primary structure (Bon et al. 2008). Functional

classification of lncRNAs includes signal, decoy, scaffold, guide, and transcriptional regulation lncRNAs based on their mode of action (Quinn and Chang 2016).

LncRNAs play important roles in development, differentiation, and disease pathogenesis (Schmitt and Chang 2016). Dysregulation of lncRNA expression has been associated with various disorders, including cancer, cardiovascular, and neurodegenerative diseases (Schmitt and Chang 2016). LncRNAs have also emerged as potential diagnostic and therapeutic targets for these diseases as I will address in Chapter 1.2.

Overall, lncRNAs are a diverse class of non-coding RNAs that have shown to play critical roles in many biological processes. A comprehensive understanding of their biological functions and cellular mechanisms will provide valuable insights into their biological roles and potential applications in disease diagnosis and treatments.

1.2 lncRNA as biomarkers for disease diagnosis and prognosis

According to many recent studies, lncRNAs are emerging as promising biomarkers for disease diagnosis and prognosis. LncRNAs are involved in multiple biological processes as I addressed in Chapter 1.1, and their dysregulation has been linked to various diseases, including cancer, cardiovascular disease, and neurodegenerative disorders (Schmitt and Chang 2016). Since some lncRNAs have their unique tissue-specific expression patterns, stability in body fluids, and differential expression in disease states, they have shown to have potential as biomarkers for the detection and monitoring of disease progression (Kunej et al. 2014).

Multiple lncRNAs have been identified as diagnostic and prognostic biomarkers for some types of cancer. For example, a lncRNA called prostate cancer antigen 3 (PCA3) has been approved as a diagnostic biomarker for prostate cancer (Leyten et al. 2015). In addition, several

lncRNAs have been shown as potential biomarkers for various types of cancer, such as HOTAIR for breast cancer and MALAT1 for lung cancer (Peng et al. 2017). Besides using lncRNAs as biomarkers to monitor cancer, in cardiovascular disease, a lncRNA called H19 has been found to be upregulated in patients with coronary artery disease and it is able to serve as a potential biomarker for diagnosis and prognosis of this disease (Wang et al. 2019). Last but not least, in neurodegenerative diseases, such as Alzheimer's disease, a lncRNA called BACE1-AS has been suggested as a potential biomarker for its diagnosis and monitoring (Fotuhi et al. 2019).

The development of lncRNA-based biomarkers is still in the early stages, but the potential applications of using lncRNAs as biomarkers in disease diagnosis and prognosis are promising. Further studies are still needed to confirm and validate the diagnostic and prognostic value of lncRNAs in clinical practice.

As suggested, lncRNAs hold great potential as biomarkers for disease diagnosis and prognosis due to their unique tissue-specific expression patterns, stability in body fluids, and differential expression in disease states (Kunej et al. 2014). LncRNA-based biomarkers have the potential to improve disease detection and monitoring, leading to earlier diagnosis and better clinical outcomes.

1.3 Role of lncRNA in cancer

Cancer is a complex disease that arises due to the accumulation of genetic and epigenetic alterations that disrupt normal cellular processes (Sadikovic et al. 2008). In recent years, a number of studies have shown that lncRNAs are involved in the development and progression of cancer. LncRNAs can be oncogene that promote cancer progression as well as tumor suppressor genes that inhibit tumor growth and metastasis (Guzel et al. 2020). Based on multiple high

throughput screening studies, lncRNAs have been shown to be deregulated in various types of cancer. Dysregulation of lncRNAs has been associated with different aspects of cancer biology, including proliferation, migration, invasion, metastasis, and drug resistance (Yousefi et al. 2020).

LncRNAs have been identified as regulators of oncogenic and tumor suppressor pathways in cancer. One well-studied lncRNA in the context of cancer is HOTAIR, which acts as a scaffold for poly-comb repressive complex 2 (PRC2) and represses the expression of target genes involved in the regulation of cellular growth and differentiation (Gupta et al. 2010). HOTAIR has been shown to be overexpressed in various types of cancer, including breast and lung cancers, and is thought to contribute to the development and progression of these cancers by promoting cellular proliferation and inhibiting apoptosis (Yu and Li, 2015). Another lncRNA that has been implicated in cancer is MALAT1, which stands for Metastasis-Associated Lung Adenocarcinoma Transcript 1. MALAT1 is highly expressed in various types of cancer, including lung and prostate cancers, and has been shown to play a role in the regulation of cell migration and invasion by modulating alternative splicing of pre-mRNA transcripts (Tripathi et al. 2013). Last but not least, a well-known lncRNA called GAS5 has been identified as a tumor suppressor in various cancers, including breast, prostate, and colon cancer, by inhibiting the activity of the oncogenic transcription factor c-Myc (Williams and Pickard 2016).

In addition to acting as oncogenes or tumor suppressors, lncRNAs can also regulate the expression of genes involved in tumor angiogenesis, a process by which tumors form new blood vessels to support their growth (Cao 2015). One example of a lncRNA involved in tumor angiogenesis is H19, which has been shown to promote angiogenesis and to be overexpressed in various types of cancer, including breast and ovarian cancers (Behera et al. 2021). Furthermore, lncRNAs have been shown to play a role in regulating cancer stem cells, which are thought to be

responsible for the initiation and maintenance of tumors (Rahman et al. 2016). For example, the lncRNA BCAR4, which stands for Breast Cancer Anti-Estrogen Resistance 4, has been shown to regulate the self-renewal and differentiation of breast cancer stem cells, suggesting that it may play a role in the development of breast cancer (Ouyang et al. 2019).

lncRNAs have been shown to be emerging as important regulators in tumorigenesis. Further research into the biological role of lncRNAs in cancer has the potential to lead to new therapeutic targets and treatments for this disease. Understanding the mechanisms by which and how lncRNAs contribute to cancer development will provide new insights into the underlying biology of this disease and could lead to the development of more effective and personalized treatments.

1.4 lncRNA databases and resources

With the increasing interest in lncRNAs, there has been a growing need for databases and resources that provide comprehensive information on lncRNA sequences, structures, functions, and interactions. Multiple lncRNA databases and resources have been developed to facilitate research on lncRNAs and to provide valuable information to the field of long non-coding RNA.

One of the most comprehensive lncRNA databases is LNCipedia, which provides information on over 150,000 human and mouse lncRNAs (Volders et al. 2013). LNCipedia provides researchers with lncRNA annotations, sequences, expression profiles, and functional information (Volders et al. 2013). Another widely used lncRNA database is NONCODE, which contains information on over 200,000 lncRNAs from various species (Zhao et al. 2016). NONCODE provides users with lncRNA sequences, structures, expression profiles, and functional annotations (Zhao et al. 2016).

In addition to these databases, several resources have been developed to predict lncRNA functions and interactions. For example, lncRNA2Target is a database that predicts lncRNA-target interactions based on sequence and expression data (Cheng et al. 2019). Another resource, lncRNome, is a platform that integrates various lncRNA-related resources, including databases, prediction tools, and experimental datasets (Bhartiya et al. 2013). Furthermore, some tools have been developed to facilitate analysis and visualization of lncRNA data. For instance, LncPipe is a comprehensive pipeline for lncRNA identification, annotation, and functional analysis (Zhao et al. 2018).

Overall, lncRNA databases and resources have become essential tools for lncRNA research, providing valuable information on lncRNA sequences, structures, functions, and interactions. These resources have helped to facilitate research on lncRNAs and to advance our understanding of the roles of lncRNAs in cellular processes and diseases.

1.5 lncRNA GRAS1

EPB41L4A-AS1, we named growth regulator antisense 1 (GRAS1), is dysregulated in multiple high throughput screening studies of different types of cancers, suggesting a possible role in regulating cancer development and progression. GRAS1 has multiple RNA variants annotated in UCSC Genome Browser, however, only part of the gene seems to be expressed based on the public RNA-seq data (Figure 1.1). GRAS1 has three exons with 1096 nts in length, as well as a snoRNA, SNORA13, in one of the introns.

GRAS1 is a long non-coding RNA that has been implicated in various types of cancer (Wang et al. 2020). It has been shown to be dysregulated in several types of cancer, including gastric cancer (Delshad et al. 2019), lung cancer (Wang et al. 2020), and colorectal cancer (Bin

et al. 2021), and is associated with tumor progression and poor prognosis (Bin et al. 2021). Regarding the functional role of GRAS1, two studies have reported that GRAS1 is involved in the regulation of cellular glycolysis and glutaminolysis through inhibiting the Warburg effect in cancer cells (Liao et al. 2019) and placental villus (Zhu et al. 2019). In the studies, they found that knockdown of GRAS1 disrupts the interaction between HDAC2 and NPM1, such that the HDAC2 can be released from nucleolus and binds on VHL and VDAC1 promoter region to promote cellular level of glycolysis and glutaminolysis (Liao et al. 2019). Another study suggests GRAS1 regulates cellular proliferation and cell cycle phase distribution by interacting with mir-146a in bone marrow-derived mesenchymal stem cells (Cui et al. 2020). However, the biological role of GRAS1 in lung cancer still remains unclear.

1.6 Lung cancer

Lung cancer is one of the leading causes of cancer-related deaths worldwide (Veena et al. 2013). It is a malignant tumor that grows in the lungs and can spread to other parts of the body. Most lung cancer cases are caused by long-term exposure to tobacco smoke (de Groot et al. 2018), but other factors such as air pollution, occupational exposure to carcinogens, and genetic predisposition can also increase the risk of developing the disease. According to a report from the World Health Organization (2020), lung cancer is the most diagnosed cancer worldwide, with an estimated 2.2 million new cases and 1.8 million deaths in 2020 alone. The survival rate for lung cancer is generally poor, with only about 19% of patients surviving for five years or more after diagnosis (National Cancer Institution, 2021). There are two main types of lung cancer: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC accounts for about 85% of all lung cancer cases, while SCLC accounts for the remaining 15%

(Mayo Clinic, 2021). Treatment options for lung cancer include surgery, radiation therapy, chemotherapy, targeted therapy, and immunotherapy (Mayo Clinic, 2021).

The etiology of lung cancer is complex, with both genetic and environmental factors contributing to its development. Smoking is the leading risk factor for lung cancer, with an estimated 80-90% of lung cancer cases attributed to tobacco smoke exposure (de Groot et al. 2018). Other risk factors include exposure to radon, asbestos, air pollution, and family history of lung cancer (Schwartz et al. 2016).

1.7 DNA damage and p53 signaling pathway

DNA damage can occur due to various environmental factors, such as ultraviolet radiation, ionizing radiation, and chemicals, as well as cellular processes, such as replication errors and oxidative stress (Barnes et al. 2018). The accumulation of DNA damage can lead to mutations, chromosomal abnormalities, and other forms of genomic instability, which can contribute to the development of cancer (Barnes et al. 2018).

The p53 tumor suppressor protein is a transcription factor that plays a crucial role in maintaining genomic stability and preventing cancer. It regulates a wide range of cellular processes, including cell cycle arrest, DNA repair, and apoptosis, to prevent the propagation of damaged DNA and to maintain the integrity of the genome (Vaddavalli and Schumacher 2022, Reinhardt and Schumacher 2012). p53 is activated in response to DNA damage and acts as a central coordinator of the cellular response to DNA damage. When DNA damage occurs in cells, p53 is activated through a series of post-translational modifications, including phosphorylation, acetylation, and ubiquitination (Reinhardt and Schumacher 2012). The activated p53 protein then translocates to the nucleus and binds to specific DNA sequences, leading to the activation or

repression of target genes. Some of the key downstream target genes of p53 signaling pathway include CDKN1A, BAX, and PIG3, which regulate the cell cycle, apoptosis, and DNA repair (Reinhardt and Schumacher 2012). P53 is considered a “guardian of the genome” because of its role in preventing the accumulation of DNA damage and the progression of cancer (Efeyan and Serrano 2007). In addition to its role in response to DNA damage, p53 also plays an important role in regulating cellular metabolism, angiogenesis, and immune surveillance (Vaddavalli and Schumacher 2022). However, despite its important role in preventing cancer, p53 is often inactivated in cancer cells, either through mutations or other oncogenic mechanisms (Selivanova and Wiman 2007). In fact, p53 mutations are among the most common genetic alterations in human cancer, and approximately 50% of all human cancers have a mutated form of p53 (Selivanova and Wiman 2007).

Overall, p53 is a crucial player in the cellular response to DNA damage, and its activation and function are essential for preventing the accumulation of DNA damage and the development of cancer. Further research into the mechanisms by which p53 is activated and regulates the cellular response to DNA damage has the potential to lead to new therapeutic targets and treatments for cancer.

1.8 RNA-antisense purification with mass spectrometry

RNA antisense purification with mass spectrometry (RAP-MS) is a technique used to study the direct interactions between RNA molecules and other cellular components, such as proteins and DNAs. The technique involves cell cross-linking using UV light, followed by purification of the RNA-interacting partners using biotinylated antisense oligonucleotide probes

that are complementary to the sequence of the RNA of interest using streptavidin beads (McHugh and Guttman 2018). The purified complexes are then analyzed by mass spectrometry, allowing for the identification of the directly interacting components (McHugh and Guttman 2018). RAP-MS is a powerful tool for investigating the function and interactions of RNA molecules in the cell. The method is highly selective, allowing for the identification of specific and direct RNA-interacting partners, and can provide a detailed understanding of the functional role of RNA in cellular processes. RAP-MS can be used to study the RNA-protein interactions that regulate gene expression, and to identify new targets for therapeutic intervention (Gottlieb and Rinn, 2016).

One well-known example of identifying proteins that directly interact with RNA using RAP-MS is XIST, which stands for X inactive specific transcript. Using this technique, it was successfully used to identify critical interacting factors of the Xist long non-coding RNA, factors that had been sought for nearly 20 years using a variety of different approaches (McHugh et al. 2015, McHugh et al. 2014). Another example of using RAP-MS is to identify proteins that are specifically and directly interacting with SARS-CoV-2 RNAs in virus-infected human cells (Schmidt et al. 2021).

Overall, RAP-MS is a valuable tool for identifying specific and direct protein interacting partners with RNA of interest that provides new insights into understanding the biological role of the RNA in cellular processes. The method is highly selective and provides a comprehensive view of the interactions between RNA molecules and proteins. RAP-MS is widely used in the study of RNA metabolism, post-transcriptional regulation, and RNA-protein interactions, and has the potential to lead to new discoveries in the field of RNA biology (Liu and Schlessinger, 2010).

1.9 Reference:

1. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C. and Kodzius, R., 2005. The transcriptional landscape of the mammalian genome. *science*, 309(5740), pp.1559-1563.
2. Liu, G., Mattick, J., & Taft, R. J. (2013). A meta-analysis of the genomic and transcriptomic composition of complex life. *Cell cycle*, 12(13), 2061-2072.
3. Wilusz, J. E., Sunwoo, H., & Spector, D. L. (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes & development*, 23(13), 1494-1504.
4. Wang, P., Xu, J., Wang, Y., & Cao, X. (2017). An interferon-independent lncRNA promotes viral replication by modulating cellular metabolism. *Science*, 358(6366), 1051-1055.
5. Zhang, L., Kang, W., Lu, X., Ma, S., Dong, L., & Zou, B. (2018). LncRNA CASC11 promoted gastric cancer cell proliferation, migration and invasion in vitro by regulating cell cycle pathway. *Cell cycle*, 17(15), 1886-1900.
6. Fatica, A., & Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Nature Reviews Genetics*, 15(1), 7-21.
7. Lingadahalli, S., Jadhao, S., Sung, Y. Y., Chen, M., Hu, L., Chen, X., & Cheung, E. (2018). Novel lncRNA LINC00844 Regulates Prostate Cancer Cell Migration and Invasion through AR Signaling Regulation of AR-Mediated Transcription by LINC00844. *Molecular Cancer Research*, 16(12), 1865-1878.
8. Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature*, 482(7385), 339-346.

9. Shi, K., Liu, T., Fu, H., Li, W., & Zheng, X. (2021). Genome-wide analysis of lncRNA stability in human. *PLoS computational biology*, 17(4), e1008918.
10. Bon, M., Vernizzi, G., Orland, H., & Zee, A. (2008). Topological classification of RNA structures. *Journal of molecular biology*, 379(4), 900-911.
11. Quinn, J. J., & Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nature Reviews Genetics*, 17(1), 47-62.
12. Schmitt, A. M., & Chang, H. Y. (2016). Long noncoding RNAs in cancer pathways. *Cancer cell*, 29(4), 452-463.
13. Kunej, T., Obsteter, J., Pogacar, Z., Horvat, S., & Calin, G. A. (2014). The decalog of long non-coding RNA involvement in cancer diagnosis and monitoring. *Critical reviews in clinical laboratory sciences*, 51(6), 344-357.
14. Leyten, G.H., Hessels, D., Smit, F.P., Jannink, S.A., de Jong, H., Melchers, W.J., Cornel, E.B., de Reijke, T.M., Vergunst, H., Kil, P. and Knipscheer, B.C., 2015. Identification of a candidate gene panel for the early diagnosis of prostate cancer. *Clinical cancer research*, 21(13), pp.3061-3070.
15. Peng, W. X., Koirala, P., & Mo, Y. Y. (2017). LncRNA-mediated regulation of cell signaling in cancer. *Oncogene*, 36(41), 5661-5667.
16. Wang, X. M., Li, X. M., Song, N., Zhai, H., Gao, X. M., & Yang, Y. N. (2019). Long non-coding RNAs H19, MALAT1 and MIAT as potential novel biomarkers for diagnosis of acute myocardial infarction. *Biomedicine & Pharmacotherapy*, 118, 109208.
17. Fotuhi, S. N., Khalaj-Kondori, M., Hoseinpour Feizi, M. A., & Talebi, M. (2019). Long non-coding RNA BACE1-AS may serve as an Alzheimer's disease blood-based biomarker. *Journal of Molecular Neuroscience*, 69, 351-359.

18. Sadikovic, B., Al-Romaih, K., Squire, J. A., & Zielenska, M. (2008). Cause and consequences of genetic and epigenetic alterations in human cancer. *Current genomics*, 9(6), 394-408.
19. Guzel, E., Okyay, T. M., Yalcinkaya, B., Karacaoglu, S., Gocmen, M., & Akçakuyu, M. H. (2020). Tumor suppressor and oncogenic role of long non-coding RNAs in cancer. *North. Clin. Istanbul*, 7, 81.
20. Yousefi, H., Maheronnaghsh, M., Molaei, F., Mashouri, L., Reza Aref, A., Momeny, M., & Alahari, S. K. (2020). Long noncoding RNAs and exosomal lncRNAs: classification, and mechanisms in breast cancer metastasis and drug resistance. *Oncogene*, 39(5), 953-974.
21. Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L. and Wang, Y., 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *nature*, 464(7291), pp.1071-1076.
22. Yu, X., & Li, Z. (2015). Long non-coding RNA HOTAIR: A novel oncogene. *Molecular medicine reports*, 12(4), 5611-5618.
23. Tripathi, V., Shen, Z., Chakraborty, A., Giri, S., Freier, S.M., Wu, X., Zhang, Y., Gorospe, M., Prasanth, S.G., Lal, A. and Prasanth, K.V., 2013. Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS genetics*, 9(3), p.e1003368.
24. Williams, G. T., & Pickard, M. R. (2016). The hormone response element mimic sequence of GAS5 lncRNA is sufficient to induce apoptosis in breast cancer cells. *Oncotarget*, 7(9).
25. Cao, Y. (2005). Tumor angiogenesis and therapy. *Biomedicine & pharmacotherapy*, 59, S340-S343.

26. Behera, J., Kumar, A., Voor, M. J., & Tyagi, N. (2021). Exosomal lncRNA-H19 promotes osteogenesis and angiogenesis through mediating Angpt1/Tie2-NO signaling in CBS-heterozygous mice. *Theranostics*, 11(16), 7715.
27. Rahman, M., Jamil, H. M., Akhtar, N., Rahman, K. M. T., Islam, R., & Asaduzzaman, S. M. (2016). Stem cell and cancer stem cell: a tale of two cells. *Progress in Stem Cell*, 3(02), 97-108.
28. Ouyang, S., Zhou, X., Chen, Z., Wang, M., Zheng, X., & Xie, M. (2019). LncRNA BCAR4, targeting to miR-665/STAT3 signaling, maintains cancer stem cells stemness and promotes tumorigenicity in colorectal cancer. *Cancer cell international*, 19(1), 1-12.
29. Volders, P.J., Helsens, K., Wang, X., Menten, B., Martens, L., Gevaert, K., Vandesompele, J. and Mestdagh, P., 2013. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic acids research*, 41(D1), pp.D246-D251.
30. Zhao, Y., Li, H., Fang, S., Kang, Y., Wu, W., Hao, Y., Li, Z., Bu, D., Sun, N., Zhang, M.Q. and Chen, R., 2016. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic acids research*, 44(D1), pp.D203-D208.
31. Cheng, L., Wang, P., Tian, R., Wang, S., Guo, Q., Luo, M., Zhou, W., Liu, G., Jiang, H. and Jiang, Q., 2019. LncRNA2Target v2. 0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic acids research*, 47(D1), pp.D140-D144.
32. Bhartiya, D., Pal, K., Ghosh, S., Kapoor, S., Jalali, S., Panwar, B., Jain, S., Sati, S., Sengupta, S., Sachidanandan, C. and Raghava, G.P.S., 2013. lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. *Database*, 2013.

33. Zhao, Q., Sun, Y., Wang, D., Zhang, H., Yu, K., Zheng, J., & Zuo, Z. (2018). LncPipe: A Nextflow-based pipeline for identification and analysis of long non-coding RNAs from RNA-Seq data. *J Genet Genomics*, 45(7), 399-401.
34. Wang, M., Zheng, S., Li, X., Ding, Y., Zhang, M., Lin, L., Xu, H., Cheng, Y., Zhang, X., Xu, H. and Li, S., 2020. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network identified lncRNA EPB41L4A-AS1 as a potential biomarker in non-small cell lung cancer. *Frontiers in Genetics*, 11, p.511676.
35. Delshad, E., Shamsabadi, F. T., Bahramian, S., Mehravar, F., Maghsoudi, H., & Shafiee, M. (2019). In silico identification of novel lncRNAs with a potential role in diagnosis of gastric cancer. *Journal of Biomolecular Structure and Dynamics*.
36. Bin, J., Nie, S., Tang, Z., Kang, A., Fu, Z., Hu, Y., Liao, Q., Xiong, W., Zhou, Y., Tang, Y. and Jiang, J., 2021. Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating the Rho/ROCK pathway in colorectal cancer. *Journal of cellular physiology*, 236(1), pp.523-535.
37. Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., Wang, Y., Wang, S., Zhu, Y., Chen, D. and Xie, W., 2019. LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. *EBioMedicine*, 41, pp.200-213.
38. Zhu, Y., Liu, Q., Liao, M., Diao, L., Wu, T., Liao, W., Wang, Z., Li, B., Zhang, S., Wang, S. and Xie, W., 2019. Overexpression of lncRNA EPB41L4A-AS1 induces metabolic reprogramming in trophoblast cells and placenta tissue of miscarriage. *Molecular Therapy-Nucleic Acids*, 18, pp.518-532.
39. Cui, P., Zhao, X., Liu, J., Chen, X., Gao, Y., Tao, K., Wang, C. and Zhang, X., 2020. miR-146a interacting with lncRNA EPB41L4A-AS1 and lncRNA SNHG7 inhibits proliferation of

bone marrow-derived mesenchymal stem cells. *Journal of Cellular Physiology*, 235(4), pp.3292-3308.

40. Veena, V. S., George, P. S., Jayasree, K., & Sujathan, K. (2013). Lung cancer is one of the leading cause of cancer related deaths for people of both sexes worldwide and sputum cytology has been accepted as one of the best non-invasive method for lung cancer diagnosis but this method is not accepted as a screening techn. *Proc. Int. J. Sci. Res. Publ.*, 2(12), 1-6.
41. World Health Organization. (2020). Cancer. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/cancer>
42. National Cancer Institute. (2021). Non-Small Cell Lung Cancer Treatment (PDQ®)–Patient Version. Retrieved from <https://www.cancer.gov/types/lung/patient/non-small-cell-lung-treatment-pdq>
43. Mayo Clinic. (2021). Lung cancer. Retrieved from <https://www.mayoclinic.org/diseases-conditions/lung-cancer/diagnosis-treatment/drc-20374629>
44. de Groot, P. M., Wu, C. C., Carter, B. W., & Munden, R. F. (2018). The epidemiology of lung cancer. *Translational lung cancer research*, 7(3), 220.
45. Schwartz, A. G., & Cote, M. L. (2016). Epidemiology of lung cancer. *Lung cancer and personalized medicine: Current knowledge and therapies*, 21-41.
46. Barnes, J. L., Zubair, M., John, K., Poirier, M. C., & Martin, F. L. (2018). Carcinogens and DNA damage. *Biochemical Society Transactions*, 46(5), 1213-1224.
47. Vaddavalli, P. L., & Schumacher, B. (2022). The p53 network: cellular and systemic DNA damage responses in cancer and aging. *Trends in Genetics*.
48. Reinhardt, H. C., & Schumacher, B. (2012). The p53 network: cellular and systemic DNA damage responses in aging and cancer. *Trends in Genetics*, 28(3), 128-136.

49. Efeyan, A., & Serrano, M. (2007). p53: guardian of the genome and policeman of the oncogenes. *Cell cycle*, 6(9), 1006-1010.
50. Selivanova, G. A., & Wiman, K. G. (2007). Reactivation of mutant p53: molecular mechanisms and therapeutic potential. *Oncogene*, 26(15), 2243-2254.
51. McHugh, C. A., & Guttman, M. (2018). RAP-MS: a method to identify proteins that interact directly with a specific RNA molecule in cells. *RNA detection: methods and protocols*, 473-488.
52. Gottlieb E, Rinn JL (2016) RNA-protein interactions in gene regulation. *Nature Reviews Genetics* 17: 703-716.
53. McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A. and Sweredoski, M.J., 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521(7551), pp.232-236.
54. McHugh, C. A., Russell, P., & Guttman, M. (2014). Methods for comprehensive experimental identification of RNA-protein interactions. *Genome biology*, 15, 1-10.
55. Schmidt, N., Lareau, C.A., Keshishian, H., Ganskih, S., Schneider, C., Hennig, T., Melanson, R., Werner, S., Wei, Y., Zimmer, M. and Ade, J., 2021. The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nature microbiology*, 6(3), pp.339-353.
56. Liu Y, Schlessinger J (2010) RNA antisense purification: a method to study RNA-protein interactions. *Nature Protocols* 5: 1064-1074.

Chapter 2: Exploring functions of novel cancer-associated long non-coding RNAs in cancer cells

2.1 Abstract

New and emerging evidence has shown the importance of lncRNA in many areas of cancer biology. However, their specific biological functions as non-coding RNA have not been well understood. In this study, we selected 5 new cancer-associated lncRNAs, including linc01977, linc00883, PDCD4-AS1, linc00909, and GRAS1 to investigate, based on screening data from multiple published studies. The pcDNA overexpression vectors of these lncRNAs were generated and validated. In addition, knockdown of the lncRNAs using the designed GapmeRs was validated as well. Using the validated overexpression and knockdown system, effects of altered expression of these lncRNAs on cell viability were tested in 5 common cell lines, including K562, HCT116, MCF7, A549, and HEK293. Interestingly, we found these lncRNAs act differently on cell viability among those cell lines. Also, we examined what transcriptomic programs were changing in the cell lines that were significantly affected on cell survival by altered expression of these lncRNAs, including overexpression and knockdown of GRAS1 in A549, as well as overexpression and knockdown of linc01977 in MCF7. Based on the PCA tests to confirm whether the biological replicates were consistent with each other, only the knockdown of GRAS1 in A549 was valid. In addition, we designed and validated biotinylated 90-mers DNA probes to capture the five lncRNA targets using RNA antisense purification with qPCR. We found three of those DNA probes, including GRAS1, linc00909, and linc00883, were able to efficiently pull down the lncRNA targets from the total RNA pool in cell lysis.

2.2 Introduction

While each human cell contains all the DNA required to create an entire organism, only some of these genes are expressed in each individual cell. Aberrant transcriptional expression (Shi et al. 2008) and epigenetic changes (Sharma et al. 2010) occur frequently in cancer cells compared to phenotypically normal cells, and although cancers are highly diverse, one common feature is their uncontrolled and inappropriate growth (Shi et al. 2008). We aim to understand how cell growth and gene expression are controlled, so that we can develop new methods and therapeutics to specifically control the growth of cancer cells compared to normal cells. Anti-cancer research has mainly focused on protein targets, and many successes in cancer treatment have been achieved with this approach. However, high-throughput sequencing studies of the human transcriptome have recently revealed the existence of a large class of ~25,000 non-coding RNA genes (Costa, 2010). These genes produce RNA transcripts that are spliced, capped, and polyadenylated in a similar manner to protein-coding genes (Statello et al. 2021), but they are not translated into proteins by the ribosome, and instead can function in their own right.

Long Non-coding RNA (lncRNA) refers to a class of non-coding RNA molecules that are greater than 200 nucleotides in length and lack the ability to encode proteins. In recent years, lncRNAs have been shown to play important roles in various biological processes, including gene regulation (Kazemzadeh et al. 2015), cellular differentiation (Yang et al. 2018), and cellular signaling pathways (Peng et al. 2017). In the context of lncRNAs and cancer, numerous studies have shown that lncRNAs can act as oncogenes, promoting cell proliferation and survival, or as tumor suppressors, suppressing cell growth and division (Wang and Chang, 2011). There are a lot of well-studied cancer-related lncRNAs, such as HOTAIR, MALAT1, GAS5, H19, and BCAR4. However, functions of a lot of lncRNAs that are connected with cancer progression

haven't been characterized, suggesting a big potential of developing new methods and therapeutics for cancer treatment.

We plan to study the functions of a selected set of lncRNAs with previously determined growth effects and correlative linkages to the development and progression of human cancers. We will analyze the protein interaction partners of five lncRNAs and determine if any transcriptional program alterations result from overexpression or knockdown of these lncRNAs in human cancer cells. We have identified five long non-coding RNAs required for cancer cell survival and that have been implicated in multiple cancer types, based on data mining of screens for cancer associated long non-coding RNAs (Yan et al. 2015, Deng et al. 2018) , published CRISPR-Cas9 screens of thousands of non-coding RNAs in cancer cell lines (Liu et al. 2017, Zhu et al. 2016) , GWAS analyses of hundreds of thousands of cancer cases and controls (Michailidou et al. 2015, Lv et al. 2017) , and individual research reports from groups studying specific cancers. Some of these lncRNAs have been studied previously, but the pathways in which they function are still unknown.

These lncRNAs are **linc01977** (Zhu et al. 2016, Michailidou et al. 2015, Guo et al. 2019, Couch et al. 2016), **linc00883** (Yan et al. 2015, Zhang et al. 2018, Zhu et al. 2016), **PDCD4-AS1** (Yan et al. 2015, Liu et al. 2017, Jadaliha et al. 2018), **linc00909** (Deng et al. 2018, Liu et al. 2017, Jadaliha et al. 2018), and **GRAS1** (Liu et al. 2017, Yang and Lv 2022, Wang et al. 2021).

For the overexpression system of these lncRNAs, Gateway cloning method was applied. Gateway Cloning is a widely used molecular biology method for efficient and rapid cloning of DNA fragments into expression vectors. It utilizes a proprietary set of recombination reactions and enzymes called "entry clones" and "destination vectors" (Invitrogen. 2011). The DNA fragment of interest is first ligated into an entry clone, which is then recombined with a

destination vector to transfer the insert into the expression vector. Gateway Cloning offers versatile and convenient options for subcloning, mutagenesis, protein expression, and downstream applications (Golic and Lindquist. 2000). Its popularity in the field of molecular biology is due to its simplicity, efficiency, and reproducibility (Sharon et al. 2003).

For the knockdown system of these lncRNAs, GapmeR knockdown method was used. GapmeR knockdown is an RNA interference (RNAi) technique used to reduce the expression of specific genes. GapmeRs are a type of DNA oligo molecule that can specifically target and degrade multiple types of RNAs, leading to a decrease in protein expression. GapmeRs are constructed using a GapmeR design software, where a 15-20 nucleotide sequence complementary to the target RNA is selected, and flanking arms are added for stability (Wang et al. 2019). The GapmeRs are then transfected into cells, where they bind to the target RNA molecules, leading to its degradation and a reduction in gene expression. GapmeR knockdown is a powerful tool for functional genomics studies and is used in various biological applications, including drug discovery and disease modeling (Li et al. 2018).

For the cell survival screening on these lncRNAs, MTT cell proliferation assay was applied. The MTT cell proliferation assay is a widely used method for measuring cellular viability and proliferation. The assay is based on the ability of living cells to convert the water-soluble yellow compound 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) into a formazan salt, which is insoluble and can be quantified spectrophotometrically (Mosmann. 1983). The MTT assay is performed by adding MTT to cells in culture, incubating for a specified time, and then measuring the amount of formazan produced, which is proportional to the number of living cells. The MTT assay is commonly used in drug discovery, toxicology studies, and to evaluate the effect of various treatments on cell proliferation (Boerma and Vermes. 2011).

For determining if any transcriptional program alterations resulting from overexpression or knockdown of these lncRNAs, bulk RNA sequencing method was applied. Bulk RNA sequencing, also known as whole transcriptome sequencing or RNA-seq, is a powerful tool for analyzing the expression of all genes in a sample at once. It provides a comprehensive view of the transcriptome, including both coding and non-coding RNAs. RNA-seq is based on high-throughput sequencing technologies that allow for the generation of millions of sequencing reads from a single sample. The first step in an RNA-seq experiment is to extract RNA from the sample, which is then converted into complementary DNA (cDNA) through reverse transcription. The cDNA is then fragmented and sequenced using either a short-read sequencing platform or a long-read sequencing platform. The sequencing data generated from RNA-seq is then aligned to a reference genome, allowing for the identification of transcribed regions and the quantification of transcript abundance. RNA-seq data can be used to identify differentially expressed genes, alternative splicing events, and non-coding RNAs. RNA-seq is widely used in many areas of biology, including gene expression analysis, functional genomics, and transcriptome-wide association studies. (Wang et al. 2009, Wang et al. 2010).

In this study, we established a system of overexpression and knockdown that is working well on our five lncRNA of interest. We assessed cell proliferation effects of the lncRNAs on 5 different types of cancer cells by MTT assay screening. In addition, we examined transcriptional program alterations resulting from overexpression or knockdown of the lncRNAs that have significant cell growth phenotype in cancer cells using bulk RNA sequencing. Also, we tested capture efficiency of the biotinylated DNA probes on the lncRNAs, which could potentially apply RNA antisense purification with Mass spectrometry technique to determine their direct-interaction protein partners and uncover their biological mechanisms in cancer cells.

2.3 Result

2.3.1 Generation of overexpression and knockdown system on lncRNA of interest

Once we confirmed our lncRNA of interest (linc01977, linc00883, PDCD4-AS1, GRAS1, and linc00909) from screening data from multiple published studies, we decided to generate pcDNA3.2 overexpression vectors that can induce ectopic expression of the lncRNA in cancer cells using the Gateway Cloning System. To make those vectors, first we cloned those lncRNA genes using PCR technique on complementary DNA (cDNA). Since some of the lncRNAs are relatively low expressed in cells, optimization in GC-content of the cloning primers and the annealing temperature for PCR reactions were required. Then we followed the protocol of the Gateway Cloning System from the manufacturer to put the insert genes into the pcDNA destination vectors. Afterwards, we sent these overexpression vectors for DNA sequencing to confirm whether they have the correct sequences. Finally, we transfected these overexpression vectors into K562 leukemia cells and validated whether they are able to induce ectopic expression via RT-qPCR. As shown in Figure 2.1, all overexpression vectors that we generated can induce ectopic expression of the lncRNAs in our cell culture system.

For the knockdown on our lncRNA of interest, we first introduced the siRNA-knockdown system, which is commonly used in protein knockdown experiments, to test whether it can still reduce expression of our target lncRNAs. We designed one siRNA that targets GRAS1 with a positive control that targets GAS5 lncRNA (target sequences of the siRNAs are listed in Figure 2.2). Using RT-qPCR to test the knockdown efficiency of the siRNAs, we found both GRAS1 and the positive control GAS5 were not able to efficiently reduce the expression of

its target RNA in K562 (Figure 2.2), suggesting that the siRNA-knockdown system does not work well on lncRNA.

Since the siRNA-knockdown system was not working well on lncRNAs, we decided to use the GapmeR-knockdown system. This knockdown system has been suggested to be applied in many lncRNA studies. We designed one GapmeR for each of our lncRNA targets through Qiagen Customized GapmeR Service (target sequences of the GapmeRs are listed in Figure 2.3). Following the similar knockdown protocol as siRNA transfection and RT-qPCR validation, we found that all the GapmeRs can significantly reduce expression in the lncRNAs (Figure 2.3).

2.3.2 Cell survival assay on lncRNA of interest

Once we validated the overexpression and knockdown systems that work efficiently on our lncRNA of interest (linc01977, linc00883, PDCD4-AS1, GRAS1, and linc00909), we decided to look into their biological functionalities. Cancer is characterized in part by a loss of cell growth control, so the question is whether altered expression of our lncRNA of interest can affect cell viability in cancer cells.

To investigate this question, we first picked five common cell lines that are able to mostly represent different types of cancer. The 5 cell lines are K562 leukemia, MCF7 breast cancer, A549 non-small lung cancer, HCT116 colon cancer, and HEK293. Then we used the established overexpression and knockdown system that we validated to change the gene expression of our lncRNA of interest in these five cell lines. Finally, we utilized MTT cell survival assay to screen and examine whether the lncRNAs have functions on controlling cell viability in these cell lines.

Overall, the results of the MTT assay screening were consistent with what a lot of studies suggest that functions of lncRNA are cell-type specific (Koch 2017). We concluded the data into

a table with individual MTT assay results in Figure 2.4. For linc01977, we found that overexpression of this lncRNA could significantly decrease cell survival in MCF7, while its knockdown led to significantly decrease cell survival in A549 and HEK293. For linc00883, we found that overexpression of this lncRNA could significantly decrease cell survival in K562 and MCF7, while its knockdown led to significantly decrease cell survival in A549, HCT116, and HEK293. For PDCD4-AS1, we found that overexpression of this lncRNA could significantly decrease cell survival in A549, HEK293 and K562, while its knockdown led to significantly decrease cell survival in A549, HEK293, and MCF7. For GRAS1, we found that overexpression of this lncRNA could significantly increase cell survival in A549 and HCT116, while its knockdown led to significantly decrease cell survival in A549 and HCT116. For linc00909, we found that knockdown of this lncRNA could significantly decrease cell survival in HCT116. Thus, these lncRNAs act differently on cell viability among those cell lines.

2.3.3 RNA sequencing on lncRNAs that affects cell viability in cancer cells

Based on the result of MTT assay screening on our lncRNA of interest, we were interested in determining if any transcriptional program alterations resulted from overexpression or knockdown of these lncRNAs that cause effects on cell viability in the cancer cells. There are three principles on selection of samples for RNA sequencing. 1) Effects of the lncRNAs on cell proliferation in the specific cancer cell type by MTT assay need to be statistically significant with $P < 0.001$ by Student's t-test. 2) RNA quality of the lncRNAs from cell extraction needs to have a RINe score larger than 9.0 (calculated by the ratio of 28S and 18S) examined by Institution for Genomic Medicine at UCSD. 3) RNA concentration of the lncRNAs from cell extraction needs to be higher than 100 ng per microliter such that it is sufficient to be run for

RNA sequencing. There were only a few samples that could pass these principles, including linc01977 overexpression in MCF7, linc01977 knockdown in MCF7, GRAS1 overexpression in A549, and GRAS1 knockdown in A549. To ensure that we are confident on our data, each group of the samples had three biological replicates. Some steps in the RNA sequencing process were completed by the Institution for Genomic Medicine at UCSD, including RNA fragmentation, cDNA generation, library preparation, and sequencing.

After receiving the raw data of the RNA sequencing, we first generated a volcano plot for each sample set, which could give us an overview of transcriptomic alternations affected by the lncRNAs. Then we ran a principal component analysis (PCA) examination on every sample set to ensure that biological replicates are consistent with each other. As shown in Figure 2.5, both overexpression and knockdown of linc01977 can significantly affect expression levels in many different genes based on the volcano plots. However, both groups of biological replicates were consistent with each other based on the PCA analysis. For the GRAS1, both overexpression and knockdown of GRAS1 can significantly affect expression levels in many different genes based on the volcano plots (Figure 2.6). Using the PCA analysis, we found that the group of GRAS1 knockdown was valid, while the group of GRAS1 overexpression had a poor PCA score (Figure 2.6). Thus, the RNAseq data of GRAS1 knockdown was valid to use for further analysis.

2.3.4 Design RAP probes to capture lncRNA of interest

Based on the result of MTT assay screening on our lncRNA of interest, we were interested in determining the protein interaction partners of our lncRNA of interest. RNA antisense purification with mass spectrometry (RAP-MS), which was used in multiple studies to sensitively identify RNA binding proteins, was applied in our study.

In order to perform the RAP-MS to identify protein interacting partners, we first designed biotinylated DNA probes that are complementary to the RNA sequence of our lncRNA targets (Figure 2.7A). To better understand the length of each biotinylated DNA probe that is required to effectively capture RNA in cell lysis, we tested capture efficiency of U1 RNA using different lengths of the U1 probes, including 60-mers, 70-mers, 90-mers, and 120-mers. In the McHugh lab, we used U1 as the positive control of our RAP experiments, because it is abundantly distributed in the cell nucleus. Our data (Figure 2.7B) shows DNA probes in both 90-mers and 120-mers could provide high capture efficiency in the elution fraction, while the probes in both 60-mers and 70-mers gave lower capture efficiency with much bigger variations. Thus, we decided to design biotinylated DNA probes for the lncRNA in 90-mers.

Once we received the biotinylated probes for every lncRNA of interest, we performed the RAP-qPCR to examine the capture efficiency on these probes in total RNA pool from cell lysis (workflow is shown in Figure 2.8). In this test, we used 90-mers U1 probes as a positive control and 90-mers luciferase probes as a non-specific negative control. Based on our data from qPCR (Figure 2.8), the probes of GRAS1, linc00883, and linc00909, were able to efficiently pull down the RNA target elution fraction without containing a significant amount of non-specific RNAs. In these experiments, we used TUBB as the non-specific RNA target.

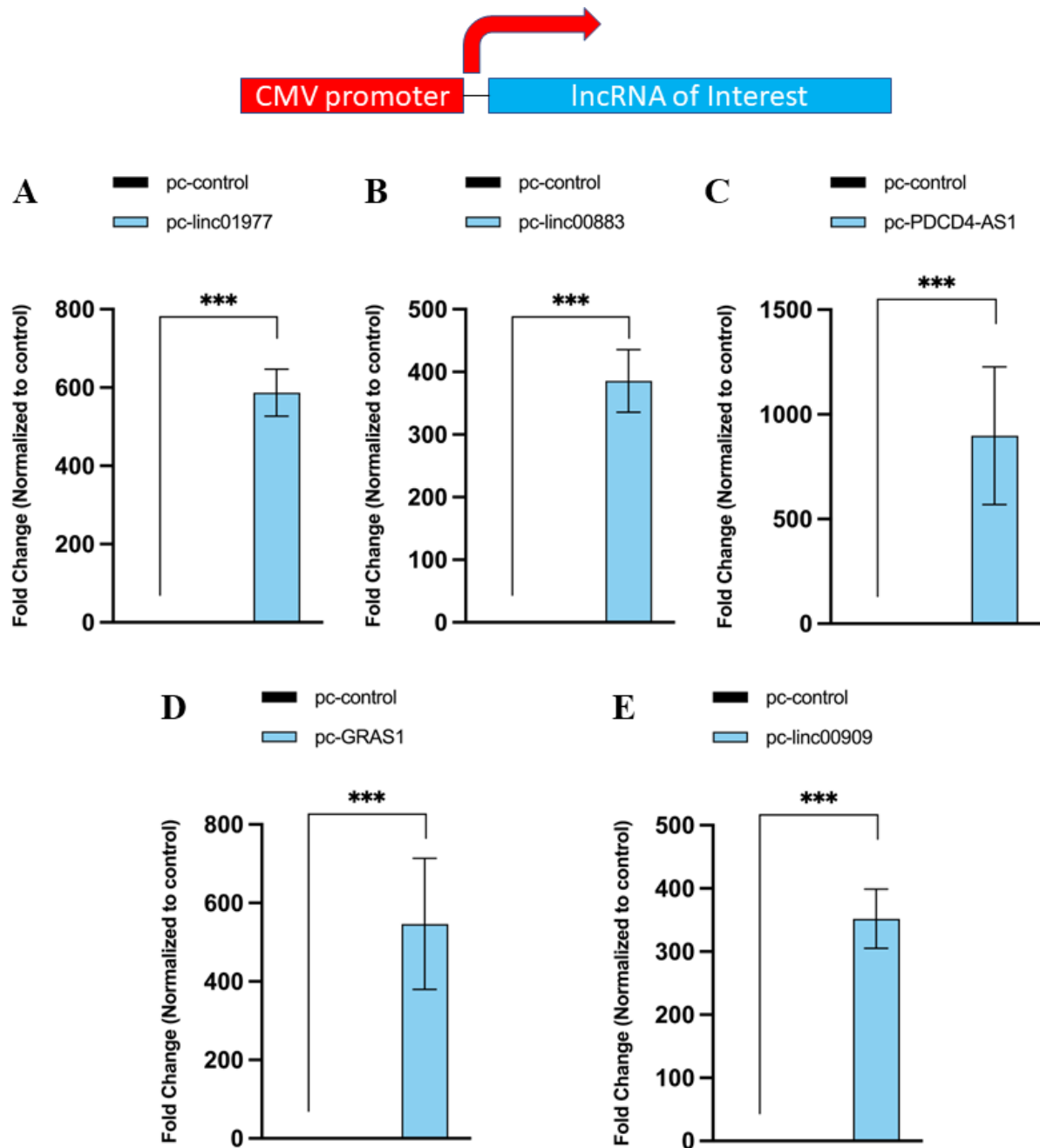


Figure 2.1: Efficient overexpression of lncRNAs in K562 Leukemia cells using pcDNA3.2 vectors. RT-qPCR of the lncRNA overexpression in K562 cells transfected with pcDNA3.2-empty vector compared with pcDNA3.2-lncRNA vectors. (A) linc01977; (B) linc00883; (C) PDCD4-AS1; (D) GRAS1; (E) linc00909. Relative fold change calculated based on the geometric mean of GAPDH RNA levels. Error bars represent SD. (***) P < 0.001 by Student's t-test.

	Target Sequence
control	UGGUUUACAUGUCGACUAA
GAS5	GAUGGAGUCUCAUGGCACA
GRAS1	GGAAAGAACUGGCCACACAUU

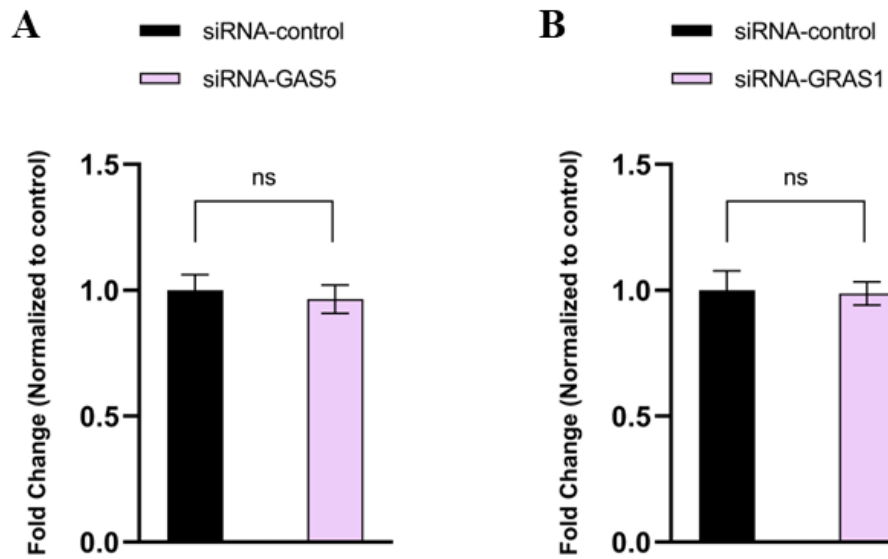


Figure 2.2: Inefficient knockdown of lncRNAs in K562 Leukemia cells using siRNA. RT-qPCR of the lncRNA knockdown in K562 cells transfected with non-targeted control siRNA compared with targeted siRNA. (A) positive control GAS5; (B) GRAS1. Relative fold change calculated based on the geometric mean of GAPDH RNA levels. Error bars represent SD. (ns) $P > 0.05$ by Student's t-test.

	Target Sequence
control	AACACGTCTATACGC
linc01977	ACATAAGAGCTTACTA
linc00883	ACGGAGCAAATCGGAA
PDCD4-AS1	GCTAGGACAGGGACTC
GRAS1	AAATGATCCAGGCAAC
linc00680	CCTGCACTTCATAACA

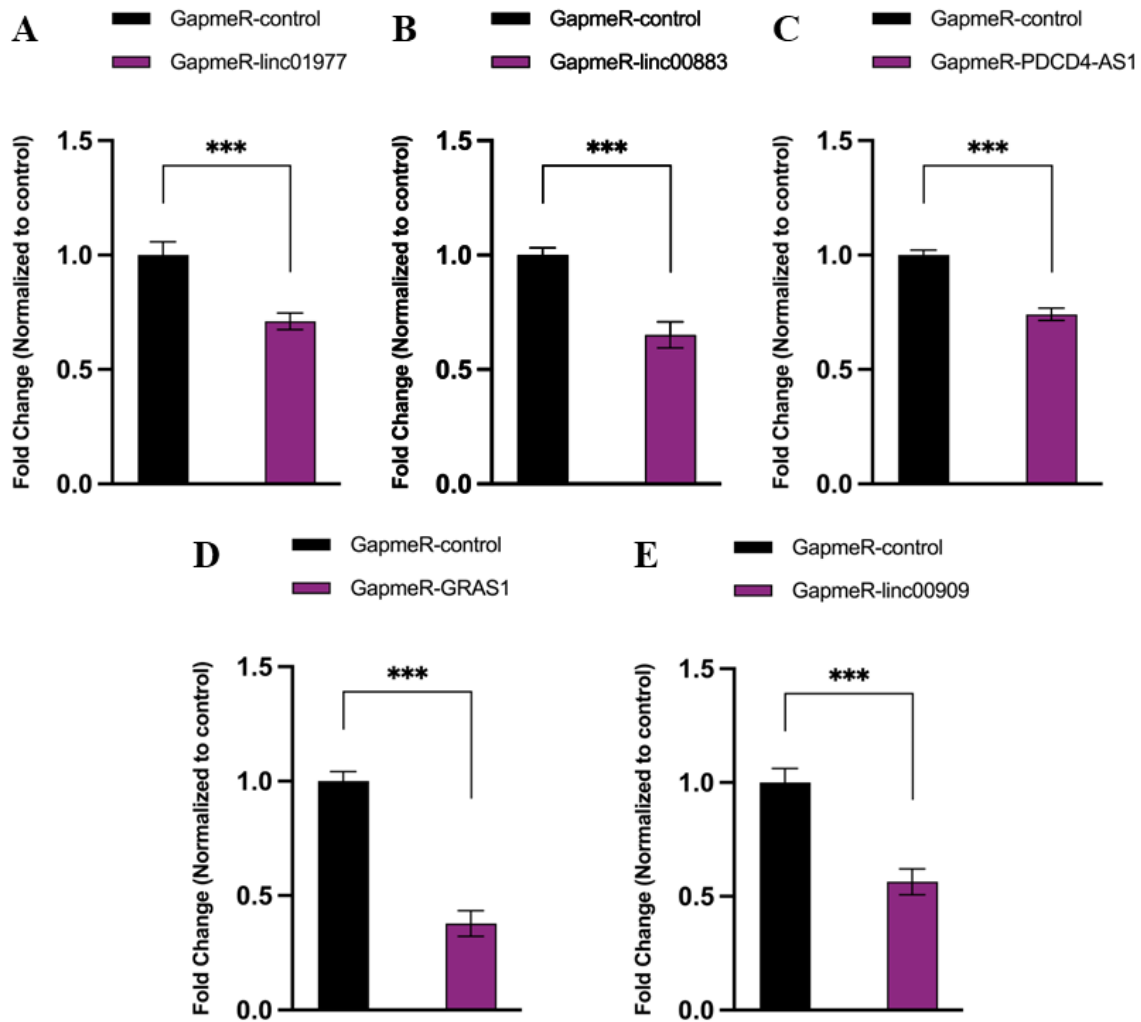


Figure 2.3: Efficient knockdown of lncRNAs in K562 Leukemia cells using GapmeRs. RT-qPCR of the lncRNA knockdown in K562 cells transfected with non-targeted control GapmeR compared with lncRNAs-targeted GapmeRs. (A) linc01977; (B) linc00883; (C) PDCD4-AS1; (D) GRAS1; (E) linc00909. Relative fold change calculated based on the geometric mean of GAPDH RNA levels. Error bars represent SD. (***) P < 0.001 by Student's t-test.

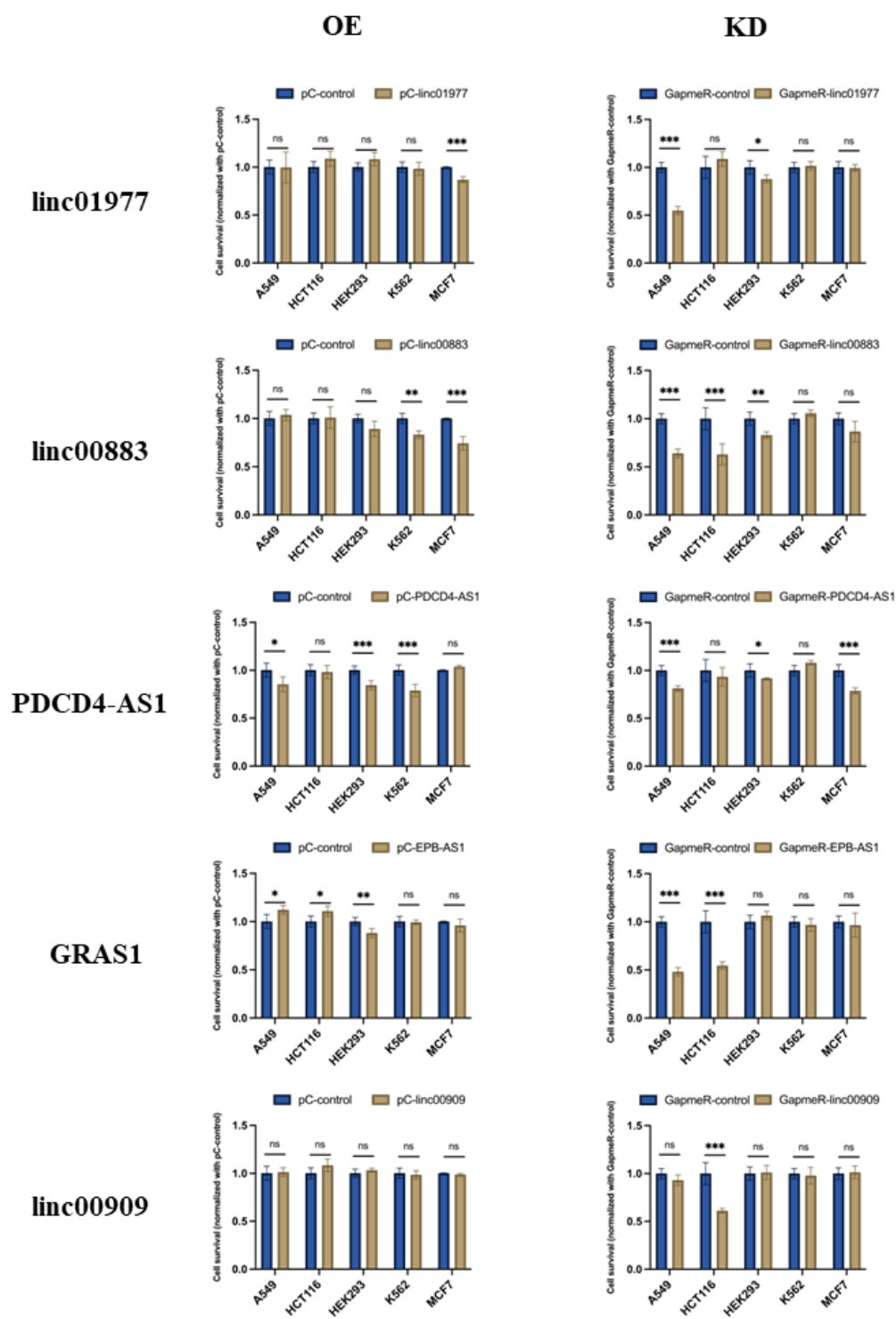


Figure 2.4 Cell survival assay on the lncRNA targets

		<u>A549</u>	<u>MCF7</u>	<u>HEK293</u>	<u>K562</u>	<u>HCT116</u>
linc00883	overexpression	NS	DOWN	NS	DOWN	NS
	knockdown	DOWN	DOWN	NS	NS	DOWN
PDCD4-AS1	overexpression	DOWN	NS	DOWN	DOWN	NS
	knockdown	DOWN	DOWN	DOWN	NS	NS
linc00909	overexpression	NS	NS	NS	NS	NS
	knockdown	NS	NS	UP	NS	DOWN
linc01977	overexpression	NS	DOWN	NS	NS	NS
	knockdown	DOWN	NS	DOWN	NS	NS
GRAS1	overexpression	UP	NS	DOWN	NS	NS
	knockdown	DOWN	NS	NS	NS	DOWN

Figure 2.4 Cell survival assay on the lncRNA targets. Cell survival was measured by MTT assay in five different cell lines, including A549, MCF7, HEK293, K562, and HCT116, transfected with either overexpression or knockdown of our lncRNA targets, which are linc00883, PDCD4-AS1, linc00909, linc01977, and GRAS1. Data are presented as fold changes normalized to each control in the bar charts. Summarized information is provided in the Table. (NS-no significant change; UP-upregulation; DOWN-downregulation). Error bars represent SD. (ns) $P > 0.05$, (***) $P < 0.001$ by Student's t-test.

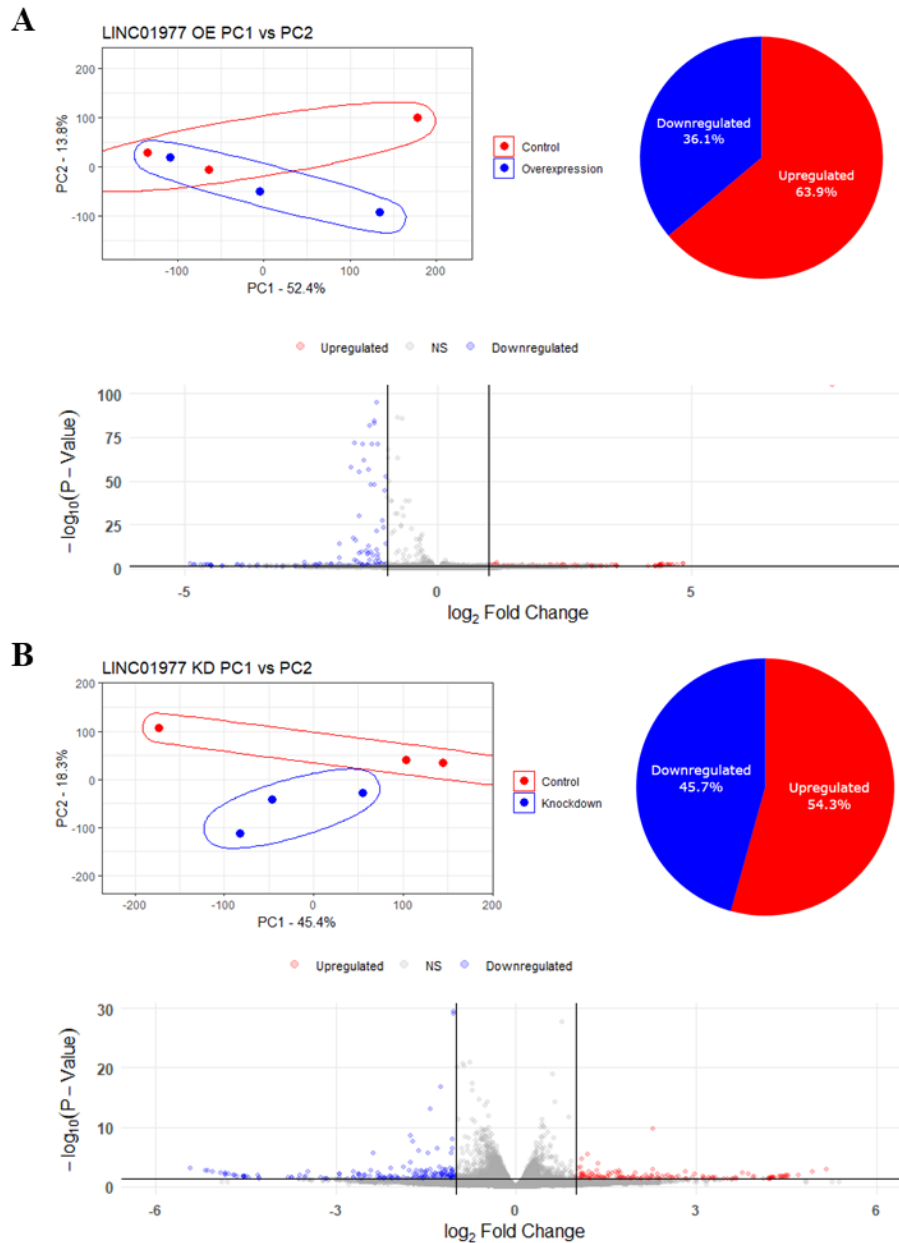


Figure 2.5 Altered expression of linc01977 affects multiple transcriptomic programs in MCF7 cells. PCA of the normalized RNAseq data transcripts per million of MCF7 transfected with control group (red) compared with linc01977 group (blue). The pie chart represents percentage of upregulation (red) and downregulation (blue) in total differential expressed genes. Volcano plot representation of differential expression analysis of genes in the control group versus linc01977 group comparison for the MCF7 cells RNAseq data sets. (A) Overexpression of linc01977. (B) knockdown of linc01977.

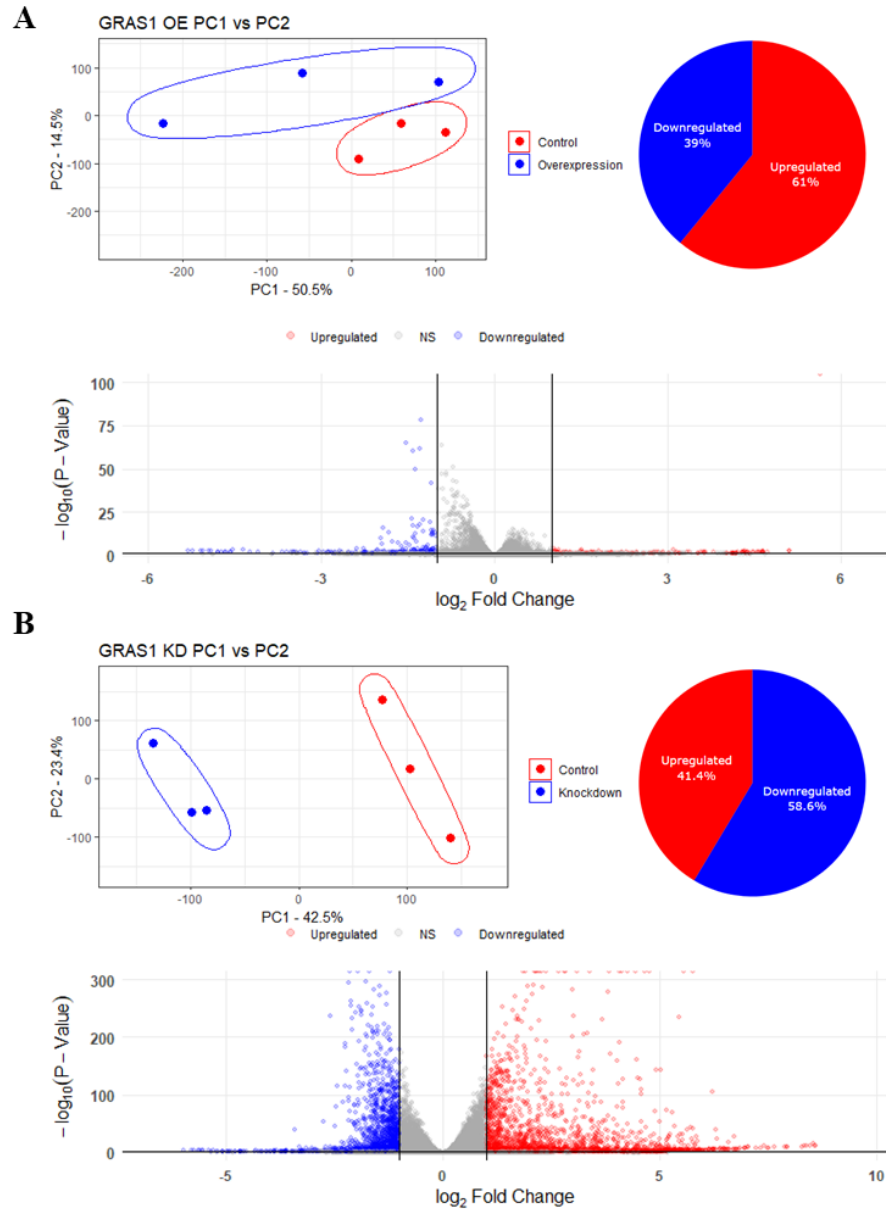
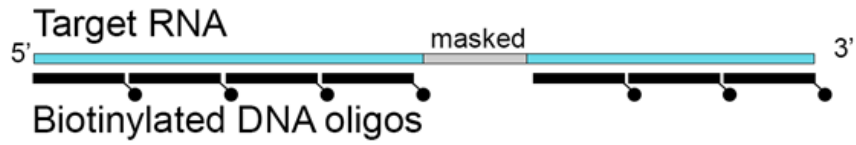


Figure 2.6 Altered expression of GRAS1 affects multiple transcriptomic programs in A549 cells. PCA of the normalized RNAseq data transcripts per million of A549 transfected with control group (red) compared with GRAS1 group (blue). The pie chart represents percentage of upregulation (red) and downregulation (blue) in total differential expressed genes. Volcano plot representation of differential expression analysis of genes in the control group versus GRAS1 group comparison for the A549 cells RNAseq data sets. (A) Overexpression of GRAS1. (B) knockdown of GRAS1.

A



B

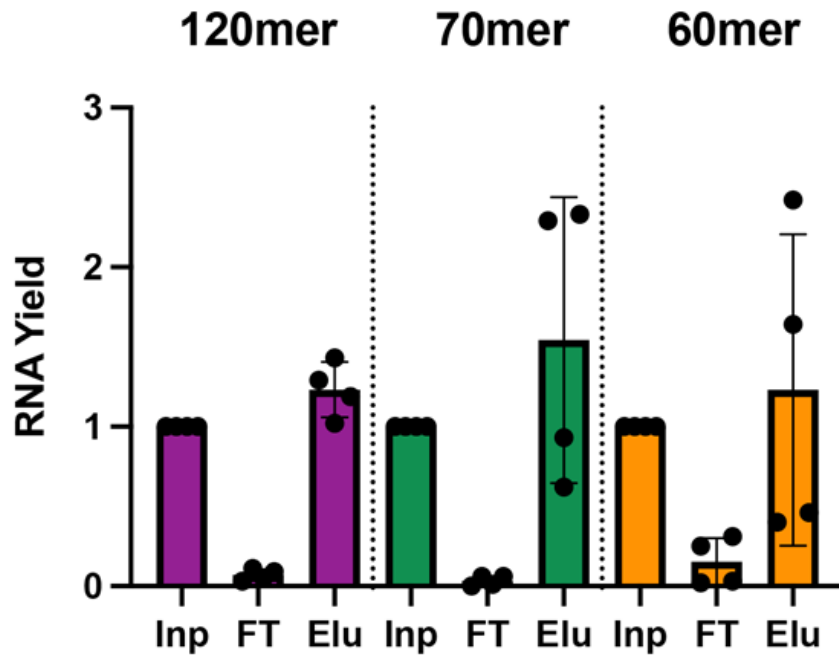


Figure 2.7 Overview of probe design and RNA target capture efficiency with different probe length. (A) Schematic design of biotinylated DNA oligos that are complementary to the target RNA sequences. DNA probes do not cover masked regions of the target RNA. (B) Four independent capture efficiency tests of U1 DNA probes that are 120mer, 90mer, 70mer, and 60mer, and recovery of the U1 RNA in the Input, Flowthrough, and Elution samples were plotted in the graph.

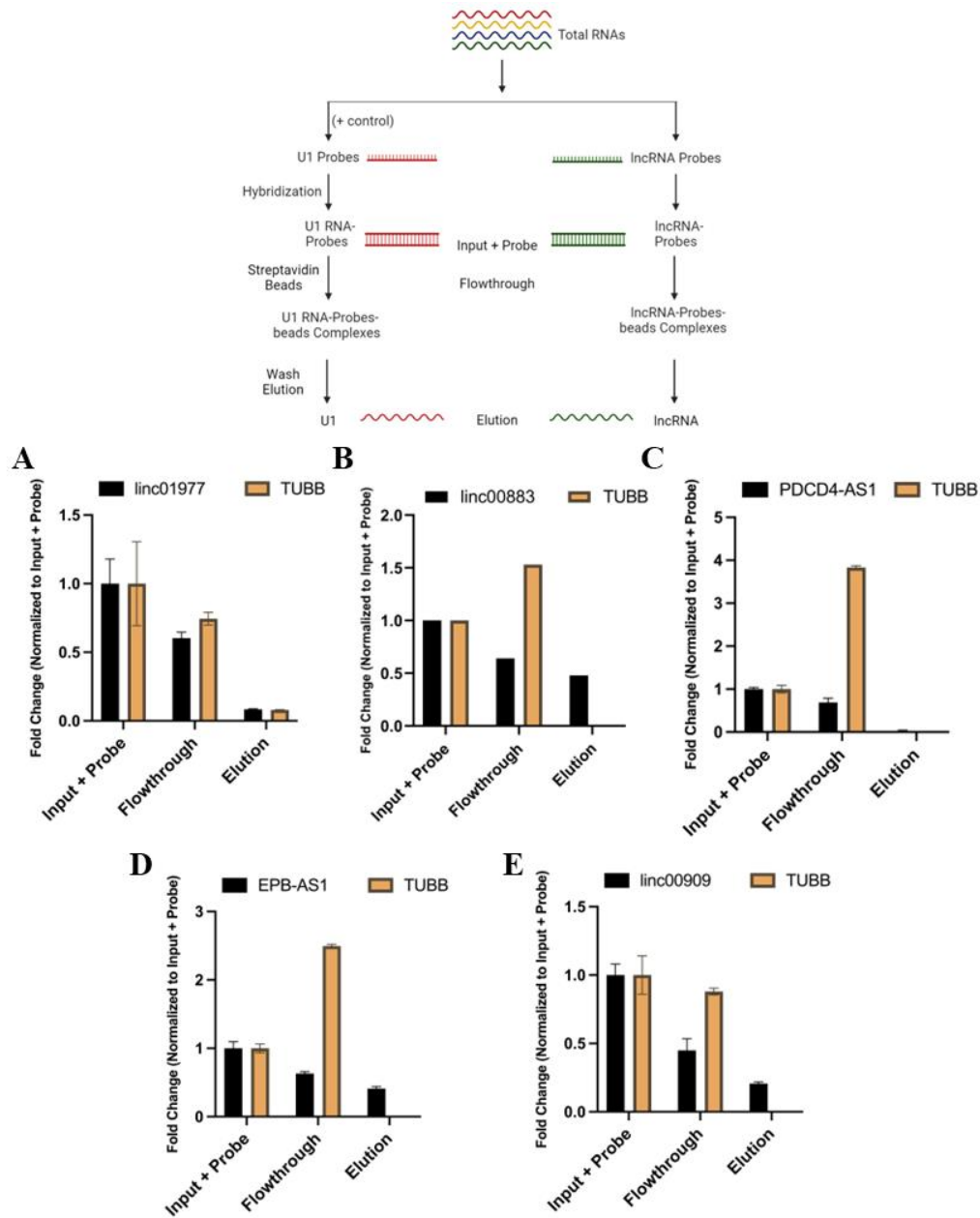


Figure 2.8 Overview of workflow on testing probe capture and lncRNA target capture efficiency. Five independent capture efficiency tests of different probes that are specific to the lncRNAs and recovery of target lncRNA and TUBB in the Input with probe, Flowthrough, and Elution samples were plotted in the graph. (A) linc01977; (B) linc00883; (C) PDCD4-AS1; (D) GRAS1; (E) linc00909.

2.4 Discussion/future directions

In this study, we picked 5 novel cancer-associated lncRNAs, including linc01977, linc00909, PDCD4-AS1, GRAS1, and linc00883, and explored their effects on cell viability in cancer cells. There are thousands of uncharacterized lncRNAs in cells, and we identified those five lncRNA of interest, based on data mining of screens for cancer associated long non-coding RNAs (Yan et al. 2015, Deng et al. 2018) , published CRISPR-Cas9 screens of thousands of non-coding RNAs in cancer cell lines (Liu et al. 2017, Zhu et al. 2016) , GWAS analyses of hundreds of thousands of cancer cases and controls (Michailidou et al. 2015, Lv et al. 2017) , and individual research reports from groups studying specific cancers. We also picked five common cell lines that were able to mostly represent different types of cancer for testing functions of these lncRNAs. The 5 cell lines are K562 leukemia, MCF7 breast cancer, A549 non-small lung cancer, HCT116 colon cancer, and HEK293.

Here, we established the overexpression and knockdown systems that were able to efficiently change expression level of the lncRNAs in our cell culture system. For the overexpression, we inserted the PCR-cloned lncRNA genes into the pcDNA overexpression vectors using Gateway Cloning system. Many of the lncRNAs have a low expression level and multiple RNA variants in cells (Tsagakis et al. 2020), including our linc00883 and GRAS1. In this study, we used PCR cloning technique to generate the lncRNA genes from cDNA, such that these genes that can be successfully cloned were predominantly expressed in the cells. However, additional work may be needed to define the functions of RNA variants of linc00883 and GRAS1 that are not well expressed in cells. For the knockdown, we found the siRNA-knockdown system worked inefficiently on reducing expression of our lncRNAs. Instead, the

GapmeR-knockdown system worked efficiently on reducing the expression of all our lncRNA targets.

From cell survival screening on altered expression of the selected lncRNAs in the five cell lines by MTT assay, we found that the lncRNAs functioned differently among these cell lines. For example, we observed that overexpression of linc00883 could significantly decrease cell survival in K562 and MCF7, while its knockdown led to significantly decreased cell survival in A549, HCT116, and HEK293. In fact, multiple studies have suggested that most lncRNAs have tissue specificity (Koch 2017, Liu et al. 2017). To gain some mechanistic insights of how these lncRNAs affect cell viability of some specific cell lines, we investigated alterations of cellular transcriptomic programs caused by different expression of the lncRNAs via RNA sequencing. Based on the volcano plot analysis using the RNAseq data, we found altered expressions of linc01977 and GRAS1 led to significant changes in expression level of many genes in MCF7 and A549 respectively. However, the PCA test, which is commonly used to validate consistency among biological replicates in the same group, showed that only knockdown of GRAS1 at A549 was valid for further analysis. Thus, repeating some of the RNA sequencing may be necessary for further study.

Several studies have shown lncRNAs can act as scaffolds to recruit functional proteins and subsequently affect biological processes (Roberts et al. 2014). Unlike regular RNA pulldown assay, RAP-MS is highly selective tool for identifying specific and direct protein interacting partners with RNA molecules, which was previously used successfully to identify the XIST non-coding RNA interaction with SPEN/SHARP in initiation of X-chromosome silencing (McHugh et al. 2015). To better understand what the lists of proteins bind to our targeted lncRNAs are, RAP-MS technique is potentially needed for the identification. In this study, we confirmed that

using biotinylated 90-mers DNA probes is able to efficiently and specifically capture RNA targets in the total RNA pool in cell lysis. We designed the DNA probes for each of our targeted lncRNAs and tested whether the probes were able to pull down the lncRNAs. We found that our probes can efficiently capture linc00883, GRAS1, and linc00909. Additional work is needed to optimize the DNA probes for linc01977 and PDCD4-AS1. Furthermore, the DNA probes for the three lncRNAs have been validated, so RAP-MS experiments for these lncRNAs to identify their protein interacting partners will be done in further study.

2.5 Reference:

1. Shi, X. B., Tepper, C. G., & deVere White, R. W. (2008). Cancerous miRNAs and their regulation. *Cell cycle*, 7(11), 1529-1538.
2. Sharma, S., Kelly, T. K., & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis*, 31(1), 27-36.
3. Costa, F. F. (2010). Non-coding RNAs: meet thy masters. *Bioessays*, 32(7), 599-608.
4. Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature reviews Molecular cell biology*, 22(2), 96-118.
5. Kazemzadeh, M., Safaralizadeh, R., & Orang, A. V. (2015). LncRNAs: emerging players in gene regulation and disease pathogenesis. *Journal of genetics*, 94, 771-784.
6. Yang, Q., Wan, Q., Zhang, L., Li, Y., Zhang, P., Li, D., Feng, C., Yi, F., Zhang, L., Ding, X. and Li, H., 2018. Analysis of LncRNA expression in cell differentiation. *RNA biology*, 15(3), pp.413-422.

7. Peng, W. X., Koirala, P., & Mo, Y. Y. (2017). LncRNA-mediated regulation of cell signaling in cancer. *Oncogene*, 36(41), 5661-5667.
8. Wang, K.C., & Chang, H.Y. (2011) "Molecular mechanisms of long noncoding RNAs." *Molecular Cell* 43(6), 904-914.
9. Yan X, Hu Z, Feng Y, Hu X, Yuan J, Zhao SD, et al. Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. *Cancer Cell*. 2015 Oct 12;28(4):529–40.
10. Deng, Y., Luo, S., Zhang, X., Zou, C., Yuan, H., Liao, G., Xu, L., Deng, C., Lan, Y., Zhao, T. and Gao, X., 2018. A pan-cancer atlas of cancer hallmark-associated candidate driver lncRNAs. *Molecular oncology*, 12(11), pp.1980-2005.
11. Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y. and Mandegar, M.A., 2017. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, 355(6320), p.eaah7111.
12. Zhu, S., Li, W., Liu, J., Chen, C.H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J. and Yuan, P., 2016. Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR–Cas9 library. *Nature biotechnology*, 34(12), pp.1279-1286.
13. Michailidou, K., Beesley, J., Lindstrom, S., Canisius, S., Dennis, J., Lush, M.J., Maranian, M.J., Bolla, M.K., Wang, Q., Shah, M. and Perkins, B.J., 2015. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nature genetics*, 47(4), pp.373-380.

14. Lv, H., Zhang, M., Shang, Z., Li, J., Zhang, S., Lian, D. and Zhang, R., 2017. Genome-wide haplotype association study identify the FGFR2 gene as a risk gene for acute myeloid leukemia. *Oncotarget*, 8(5), p.7891.
15. Golic KG, Lindquist S (2000) A Gateway to the Function of Uncharacterized Genes. *Science* 287: 1920-1921.
16. Invitrogen (2011) Gateway Cloning Technology. Thermo Fisher Scientific.
<https://www.thermofisher.com/us/en/home/life-science/cloning/gateway-cloning-technology.html>
17. Sharon D, Muller K, van der Meer AD (2003) The Gateway® technology: a versatile platform for recombinant protein production. *Trends in Biotechnology* 21: 123-127.
18. Li Y, Fan L, Shao N (2018) RNA interference technology and its application in cancer research. *Oncology Letters* 15: 3765-3772.
19. Wang W, Fan J, Liu Z (2019) GapmeR knockdown of long non-coding RNA HOTAIR affects the proliferation and apoptosis of colorectal cancer cells. *Oncology Letters* 17: 5155-5162.
20. Boerma M, Vermes I (2011) The MTT assay in cellular pharmacology. *Journal of Molecular and Cellular Cardiology* 51: 674-680.
21. Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of Immunological Methods* 65: 55-63.
22. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10: 57-63.
23. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. and Chen, Z., 2011. Full-length transcriptome

- assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), pp.644-652.
24. Guo, K., Chen, L., Wang, Y., Qian, K., Zheng, X., Sun, W., Sun, T., Wu, Y. and Wang, Z., 2019. Long noncoding RNA RP11-547D24. 1 regulates proliferation and migration in papillary thyroid carcinoma: Identification and validation of a novel long noncoding RNA through integrated analysis of TCGA database. *Cancer medicine*, 8(6), pp.3105-3119.
 25. Couch, F.J., Kuchenbaecker, K.B., Michailidou, K., Mendoza-Fandino, G.A., Nord, S., Lilyquist, J., Olswold, C., Hallberg, E., Agata, S., Ahsan, H. and Aittomäki, K., 2016. Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nature communications*, 7(1), p.11375.
 26. Yan, X., Hu, Z., Feng, Y., Hu, X., Yuan, J., Zhao, S.D., Zhang, Y., Yang, L., Shan, W., He, Q. and Fan, L., 2015. Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer cell*, 28(4), pp.529-540.
 27. Zhang, G., Pian, C., Chen, Z., Zhang, J., Xu, M., Zhang, L. and Chen, Y., 2018. Identification of cancer-related miRNA-lncRNA biomarkers using a basic miRNA-lncRNA network. *PloS one*, 13(5), p.e0196681.
 28. Jadaliha, M., Gholamalamdari, O., Tang, W., Zhang, Y., Petracovici, A., Hao, Q., Tariq, A., Kim, T.G., Holton, S.E., Singh, D.K. and Li, X.L., 2018. A natural antisense lncRNA controls breast cancer progression by promoting tumor suppressor gene mRNA stability. *PLoS genetics*, 14(11), p.e1007802.
 29. Yang, F., & Lv, S. (2022). LncRNA EPB41L4A-AS1 Regulates Cell Proliferation, Apoptosis and Metastasis in Breast Cancer. *Annals of Clinical & Laboratory Science*, 52(1), 3-11.

30. Wang, Z., Liao, W., Liu, F., Yang, T., Xie, W., Liao, M., Gu, D. and Zhang, Y., 2021. Downregulation of lncRNA EPB41L4A-AS1 mediates activation of MYD88-dependent NF- κ B pathway in diabetes-related inflammation. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, pp.265-277.
- Koch, L. (2017). Screening for lncRNA function. *Nature Reviews Genetics*, 18(2), 70-70.
31. Tsagakis, I., Douka, K., Birds, I., & Aspden, J. L. (2020). Long non-coding RNAs in development and disease: conservation to mechanisms. *The Journal of pathology*, 250(5), 480-495.
32. Roberts, T. C., Morris, K. V., & Weinberg, M. S. (2014). Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*, 9(1), 13-20.
33. McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A. and Sweredoski, M.J., 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521(7551), pp.232-236.

2.6 Acknowledgement

Chapter 2, in full, is currently being prepared for publication of the material. The authors in the publication: Tong Su, Calvin Huang, Lingbo Kong, Jonathan Zhu, and Colleen McHugh. The dissertation author is the primary investigator and author of this material.

CHAPTER 3: GRAS1 long non-coding RNA directly interacts with STAT5 to regulate STAT5 activity

3.1 Abstract

Long non-coding RNAs have been shown to be involved in multiple biological processes. However, their specific biological functions as non-coding RNA are not well understood. In this study, we aimed to identify the direct interacting protein partners with a lncRNA, termed Growth Regulator Antisense 1 (GRAS1). Using RNA-antisense purification with mass spectrometry (RAP-MS), we identified signal transducer and activator of transcription 5 (STAT5) directly and specifically binds to GRAS1 in K562 cells. While altered expression of GRAS1 does not affect cell survival in the cell line, knockdown of the GRAS1 reduces phosphorylation level of the STAT5 as well as the expression of STAT5 downstream target genes. In addition, we found U1 RNA binds with similar interacting proteins in different cell lines using the RAP-MS.

3.2 Introduction

Long non-coding RNAs (lncRNAs) are a group of non-coding RNAs that are longer than 200 nucleotides and do not encode proteins. Despite lacking protein-coding capabilities, lncRNAs have emerged as key regulators of gene expression and can affect various biological processes, including proliferation, differentiation, and disease development (Yoon et al. 2013, Bhan and Mandal 2014, Huarte 2015). LncRNAs are involved in many different biological processes, such as epigenetic regulation, transcriptional and post-transcriptional regulation, and signaling pathways (Huarte 2015, He et al. 2019, Chen et al. 2018). LncRNAs have been shown to regulate gene expression by various mechanisms, including chromatin remodeling,

transcriptional regulation, and post-transcriptional regulation (Huarte 2015, Yoon et al. 2013). For example, the lncRNA HOTAIR has been found to regulate the expression of multiple genes by recruiting chromatin-modifying enzymes to specific genomic loci, leading to changes in chromatin structure and gene expression (Holoch and Moazed 2015, Koziol and Rinn 2010). Further research into the mechanisms of lncRNA function and their potential as diagnostic and therapeutic targets may lead to new insights into the regulation of gene expression and the development of new treatments for diseases.

GRAS1 is a long non-coding RNA that has been implicated in various types of cancer (Wang et al. 2020). It has been shown to be dysregulated in several types of cancer, including gastric cancer (Delshad et al. 2019), lung cancer (Wang et al. 2020), and colorectal cancer (Bin et al. 2021), and is associated with tumor progression and poor prognosis (Bin et al. 2021). Regarding the functional role of GRAS1, two studies have reported that GRAS1 is involved in the regulation of cellular glycolysis and glutaminolysis through inhibiting the Warburg effect in cancer cells (Liao et al. 2019) and placental villus (Zhu et al. 2019). In the studies, they found that knockdown of GRAS1 disrupts the interaction between HDAC2 and NPM1, such that the HDAC2 can be released from nucleolus and binds on VHL and VDAC1 promoter region to promote cellular level of glycolysis and glutaminolysis (Liao et al. 2019). Another study suggests GRAS1 regulates cellular proliferation and cell cycle phase distribution by interacting with mir-146a in bone marrow-derived mesenchymal stem cells (Cui et al. 2020).

A number of studies have shown lncRNAs can act as scaffolds to recruit functional proteins and subsequently affect biological processes (Roberts et al. 2014). Unlike regular RNA pulldown assay, RAP-MS is highly selective tool for identifying specific and direct protein interacting partners with RNA molecules, which was previously used successfully to identify the

XIST non-coding RNA interaction with SPEN/SHARP in initiation of X-chromosome silencing (McHugh et al. 2015). The technique involves cell cross-linking using UV light, followed by purification of the RNA-interacting partners using biotinylated antisense oligonucleotide probes that are complementary to the sequence of the RNA of interest using streptavidin beads (McHugh and Guttman 2018). The purified complexes are then analyzed by mass spectrometry, allowing for the identification of the directly interacting components (McHugh and Guttman 2018). RAP-MS can be used to study the RNA-protein interactions that regulate gene expression, and to identify new targets for therapeutic intervention (Gottlieb and Rinn, 2016).

Signal transducer and activator of transcription 5 (STAT5) is a member of the STAT family of transcription factors, which play crucial roles in various cellular processes such as cell proliferation, differentiation, and survival (Calò et al. 2003, Coffey et al. 2000, de Groot et al. 1998). Several studies have shown that STAT5 is involved in various physiological and pathological processes, such as immune response, hematopoiesis, mammary gland development, and cancer (Calò et al. 2003, Buettner et al. 2002, Halim et al. 2020). Aberrant activation of STAT5 has been implicated in the pathogenesis of several types of cancer, including leukemia, breast cancer, and prostate cancer (Calò et al. 2003, Buettner et al. 2002). STAT5 is activated by cytokines, growth factors, and hormones that bind to their corresponding receptors on the cell surface, leading to the phosphorylation of STAT5 by Janus kinases (Coffey et al. 2000, de Groot et al. 1998, Rane and Reddy 2000). The phosphorylated STAT5 forms homodimers or heterodimers and translocates to the nucleus, where it binds to specific DNA sequences, and activates or represses the expression of target genes (Cella et al. 1998, Paukku and Silvennoinen 2004).

In this study, we identified GRAS1 long non-coding RNA directly interacts with STAT5 and regulates STAT5 phosphorylation at K562 cells. Downregulation of GRAS1 by GapmeR knockdown reduced phosphorylation of STAT5 as well as downstream target genes of STAT5 pathway. Altered expression of GRAS1 had no significant effects on cell survival at K562 cells. In addition, we found GRAS1 is mostly located in the nucleus by subcellular fractionation at K562.

3.3 Result

3.3.1 U1 binds to similar protein interacting partners in different cell lines

Based on the MTT assay screening on GRAS1, we found that overexpression of GRAS1 increased cell survival, while knockdown of GRAS1 dramatically decreased the cell survival in A549 cells, as shown in Figure 2.4. As we addressed in Chapter 2, we successfully completed the RNA sequencing on the GRAS1 knockdown samples with a good PCA score (Figure 2.7). Using RAP-qPCR technique, we validated the 90-mers DNA probes that were able to effectively capture GRAS1 in total RNA pool from cell lysis along with the U1 positive control in the experiment (Figure 3.1). Thus, we decided to perform RNA antisense purification with mass spectrometry (RAP-MS) to determine protein interacting partners with the GRAS1 RNA.

In order to get a list of proteins with a good MS intensity, a sufficient amount of cells are required as the input. We first tested RAP-MS on U1 RNA using 50M K562 cells, because proteins that directly bind to U1 are known (So et al, 2019). RAP-MS experimental scheme is shown in Figure 3.2A. As shown in Figure 3.2C, the known U1 interacting factors were of highest intensity from the RAP-MS experiment, including multiple small nuclear ribonucleoproteins (SNRPG, SNRNP70, SF3B1, SNRPD, SNRPA) and heterogeneous nuclear

ribonucleoproteins (HNRNPC, HNRNPA, HNRNPM, HNRNPU, HNRNPR). The result of this RAP-MS on U1 suggests 50M cells is sufficient to get a good output from the mass spectrometry without the use of SILAC or other specific protein labeling methods.

Although the effect of GRAS1 on cell survival was only seen in A549, K562 cells were preferred to be used instead to perform the RAP-MS experiments for two reasons: 1) K562 is a suspension cell line with a much higher proliferation rate compared with A549. It is much easier to acquire a large amount of K562 cells from a cell bucket (Figure 3.3A) than A549 cells from a stack of treated flasks. 2) We hypothesized that protein interacting partners with GRAS1 are consistent in the two cell lines.

To test the hypothesis that lncRNAs should bind to similar protein interacting partners in different cell lines, we performed another RAP-MS on U1 RNA in both A549 and HCT116 for comparison with the protein list from RAP-MS in K562. As shown in Figure 3.2B-3.2D, top protein interacting factors of highest MS intensity were still mostly nuclear ribonucleoproteins in the protein list from the RAP-MS in these three cell lines, suggesting U1 interacts with similar proteins in different cells. Thus, we decided to use K562 cells to perform the RAP-MS on the GRAS1.

3.3.2 STAT5 directly interacts with GRAS1 at K562

Since GRAS1 is much less abundant compared with the U1 in cells, we decided to increase the cell input to be 150M for the GRAS1 RNA, along with 50M for U1 as the positive control and 50M for luciferase as the negative control, per RAP-MS experiment.

To get better confidence to identify the direct-interacting proteins with GRAS1, we did the RAP-MS experiments two times. For our MS analysis, we filtered out any background

proteins found in the luciferase group. To increase the specificity of GRAS1 interacting proteins, we also removed proteins identified in the U1 group, because GRAS1 functions differently compared with U1 in cells. From the first RAP-MS experiment, 95 proteins with a minimum of 2 unique peptides were pulled down and identified in the GRAS1 group. From the second RAP-MS experiment, we found a list of 116 proteins with a minimum of 2 unique peptides in the GRAS1 group. Only 11 proteins were identified in both RAP-MS experiments, and they were chosen as candidates for GRAS1 interactors (Table). Among these proteins, we found multiple mitochondrial proteins that are involved in cellular metabolism (HSPA9, MYH9, ATP5B, DYNC1I2, SMARCA5), two proteins that link to chromosome regulation (SMC4, IMMT), and one proteasome-related protein (PSMD6). STAT5, not only was of the highest MS intensity among these 11 identified GRAS1 interactors, but it is also suggested to have biological functions in cell viability in multiple studies (Harir et al. 2008, Matsumura et al. 1999). Thus, we decided to investigate whether the interaction between GRAS1 and STAT5 has biological significance.

To validate the interaction between STAT5 and GRAS1 from the RAP-MS, we performed RNA immunoprecipitation using the K562 cell lysis. To ensure the antibody was able to pull down sufficient amounts of endogenous STAT5 at K562, we did an immunoblotting to probe for STAT5 after the pulldown. As shown in Figure, STAT5 was efficiently pulled down using the antibody against STAT5, while IgG control had no capture of STAT5. Afterwards, an approximately 20-fold enrichment of GRAS1 was observed in the STAT5 pulldown samples compared with IgG control samples via RT-qPCR (Figure), suggesting that STAT5 protein interacts with GRAS1 in K562 cells.

To confirm the interaction is still valid at A549, we repeat the RNA immunoprecipitation on A549 cell lysis. Surprisingly, no significant enrichment of GRAS1 was observed in the STAT5 pulldown samples compared with IgG control samples (Figure). That may be caused by low cellular concentration of STAT5 at A549 compared with K562. To test this hypothesis, immunoblotting was used to probe for STAT5 on both A549 and K562 cell lysis. As shown in Figure, STAT5 was clearly shown at K562 cell lysis while it was not detected at A549 cell lysis, suggesting that K562 cells have much higher cellular amount of STAT5 protein than A549 cells. Consistently, no endogenous STAT5 was observed in the STAT5 pulldown using A549 cell lysis (Figure). In fact, many studies demonstrate that STAT5 has biological functions in the immune system (Verdeil et al. 2019, Rani and Murphy. 2016), leukemia (Shuai et al. 1996, Warsch et al. 2011), and breast cancer (Walker et al, 2009), while few studies suggest that STAT5 links with lung cancer. In conclusion, the effects of GRAS1 on cell viability that we observed in A549 from the MTT assay screening have no connection with the STAT5 protein.

3.3.3 GRAS1 does not affect cell survival at K562

While STAT5 is not the protein target related to the cell survival in A549, it is critical for cancer progression in Leukemia (Shuai et al, 1996). We decided to investigate whether the interaction between STAT5 and GRAS1 has cellular functions in K562.

Based on the MTT screening on GRAS1, no significant effect on cell survival was observed in K562 cells when GRAS1 was upregulated or downregulated (Figure 2.4). To validate the MTT assay, a more sensitive cell viability detection kit, CCK6, was applied. Consistent with the MTT assay, we found no significant difference in cell survival in either GRAS1 overexpression or GRAS1 knockdown at K562 (Figure). Also, annexin V - PI apoptosis

assay was used to examine the cell apoptosis at K562. As shown in Figure, no significant change in overall apoptotic population, including the early (Annexin V+/PI-) and late (Annexin V+/PI+) apoptotic cells, when the GRAS1 got upregulated or downregulated in K562 cells. Thus, GRAS1 does not affect cell viability in K562 cells.

3.3.4 Knockdown of GRAS1 affects phosphorylation of STAT5

STAT5 is shown to bind DNA as a transcription factor (Stoecklin et al, 1997), but few studies suggest it is also an RNA-binding protein. Based on the RAP-MS and RNA immunoprecipitation, we confirmed that STAT5 is able to bind with GRAS1 in K562 cells. We decided to investigate whether GRAS1 has regulatory effects on STAT5 through the interaction. As shown in Figure, both overexpression of GRAS1 and knockdown of GRAS1 had no effect on the total protein level of STAT5. Interestingly, we found the knockdown of GRAS1 significantly decreased the phosphorylation level of STAT5 (Figure), while no significant difference in the STAT5 phosphorylation was observed when GRAS1 got upregulated in K562 (Figure).

3.3.5 Knockdown of GRAS1 affects subcellular localization of STAT5

Phosphorylation is important for STAT5 translocation from cytoplasm to nucleus such that STAT5 can act as a transcription factor to activate multiple oncogenic expressions. (Schmidt et al, 2014). As shown in Figure, knockdown of GRAS1 significantly decreased typical downstream gene targets of STAT5 signaling pathway, including p85, p110, and AKT1. However, overexpression of GRAS1 had no significant effects on these downstream genes (Figure). To better understand whether the effect on phosphorylation level of STAT5 caused by GRAS1 knockdown leads to an increase in nuclear composition of STAT5 at K562, fluorescent

imaging was applied. A plasmid that can make STAT5 with an eGFP tag fusion protein (eGFP-STAT5) was made via PCR cloning. Using imaging under GFP channel and immunoblotting to probe for GFP, eGFP-STAT5 was validated to be expressed at K562 transfected with the plasmid (Figure). As shown in Figure, knockdown of GRAS1 had a significant decrease in the intensity of eGFP signal within the DAPI-stained region compared with the control, suggesting a decrease in the nuclear level of STAT5 when GRAS1 got downregulated in K562.

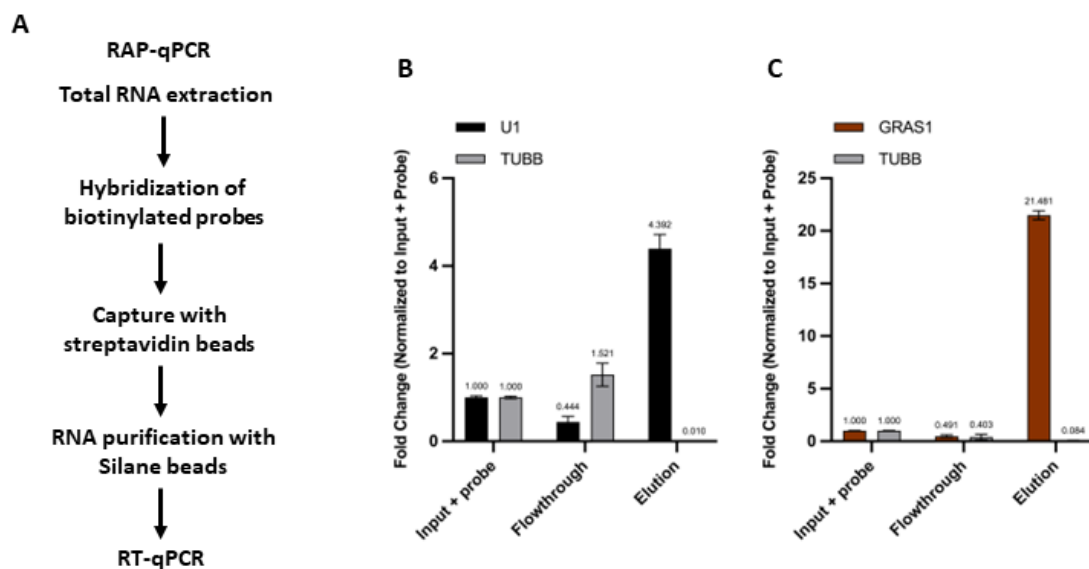


Figure 3.1 Overview of workflow on RAP-qPCR and testing capture efficiency of the DNA probes. Two independent capture efficiency tests of different probes that are specific to the RNA and recovery of RNA target and TUBB in the Input with probe, Flowthrough, and Elution samples were plotted in the graph. (A) RAP-qPCR experimental scheme to validate biotinylated DNA probes. (B) U1 capture using the U1 probes. (C) GRAS1 capture using the GRAS1 probes.

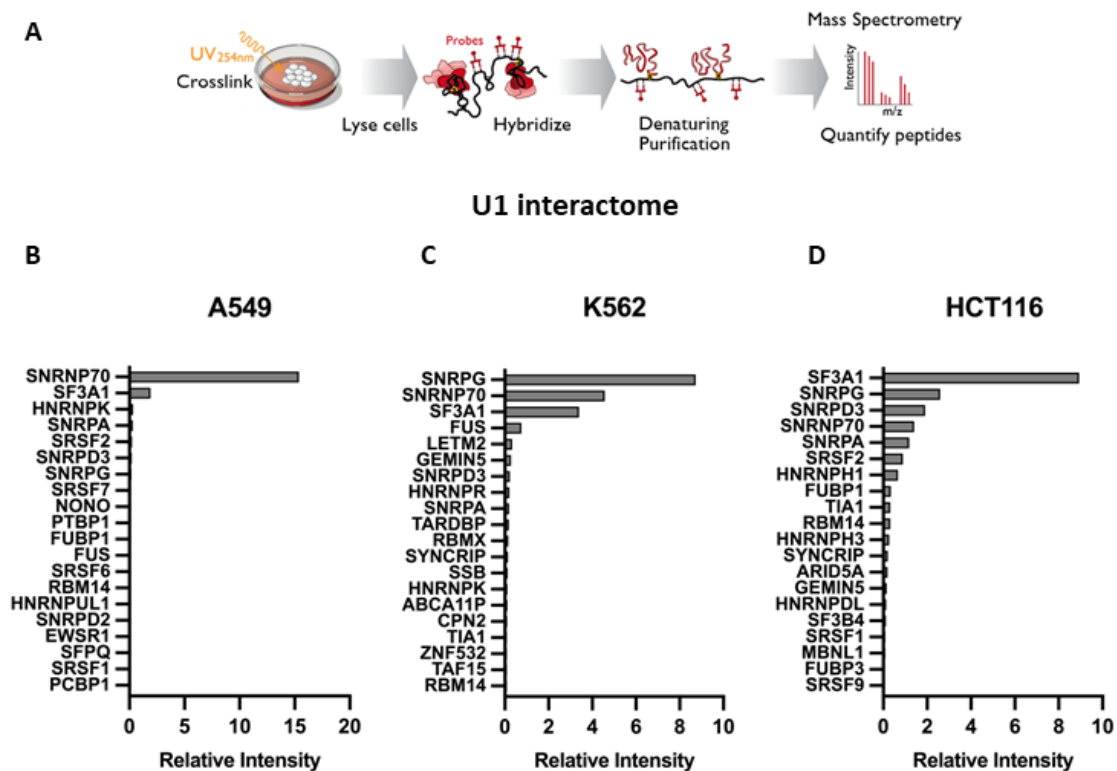
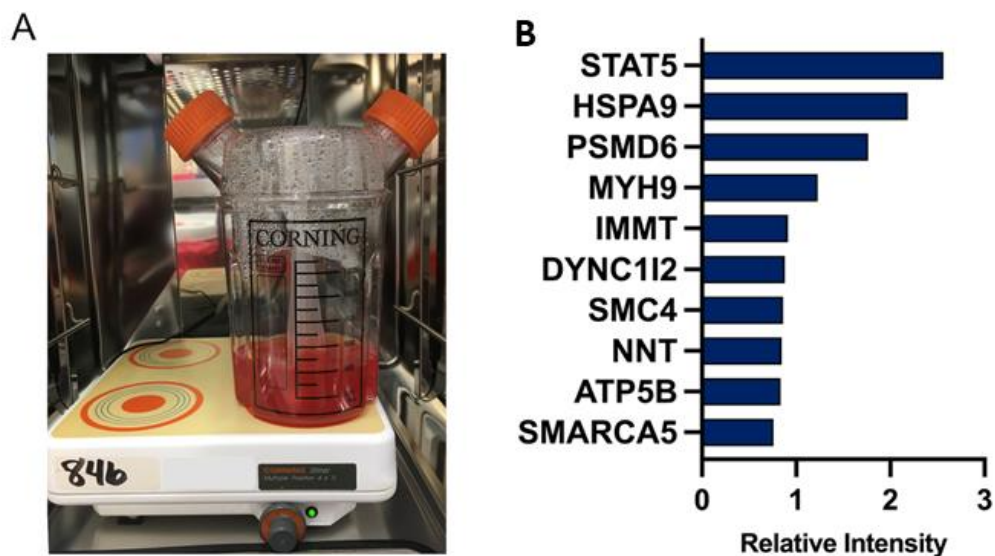


Figure 3.2 U1 binds to similar protein interacting partners in different cell lines. 20 U1-interacting proteins with highest MS intensity, excluding proteins found in luciferase samples, are shown in the bar charts. (A) RNA antisense purification with mass spectrometry experimental scheme on U1. (B) Top 20 U1 interacting proteins in A549. (C) Top 20 U1 interacting proteins in K562. (D) Top 20 U1 interacting proteins in HCT116.



Protein name	Unique peptides (MS1)	Unique peptides (MS2)	Molecular weight (KDa)
STAT5	2	5	90.646
HSPA9	5	4	73.68
PSMD6	4	3	45.531
MYH9	2	6	226.53
IMMT	3	4	83.677
DYNC1I2	3	3	71.456
SMC4	2	5	147.18
NNT	3	2	113.89
ATP5B	2	2	56.559
SMARCA5	2	3	121.9

Figure 3.3. GRAS1 directly interacts with STAT5 in K562 cells. (A) Experimental condition to grow K562 cells in a cell bucket. (B) 10 GRAS1-interacting proteins with highest MS intensity, excluding proteins found in RAP-MS of U1 and luciferase samples (top). Table of information on these 10 proteins with unique peptides and molecular weight (bottom).

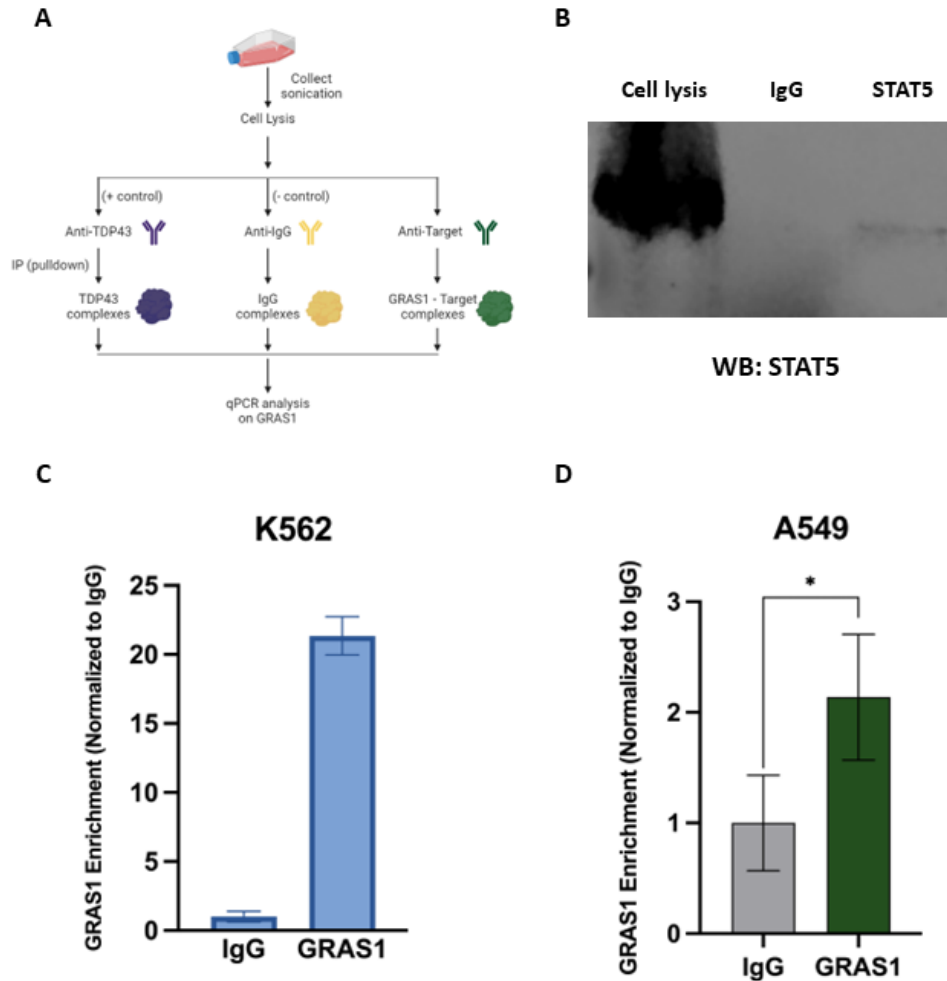


Figure 3.4 GRAS1 was confirmed to bind to STAT5 in K562 cells, but the interaction is not valid in A549 cells. (A) Experimental scheme of RNA immunoprecipitation. (B) Immunoblotting probed for STAT5 after STAT5 was pulldown using anti-STAT5 in K562 cell lysis. IgG is a negative control in the experiment. (C) GRAS1 enrichment measured by RT-qPCR after immunoprecipitation on STAT5 in K562 cell lysis. (D) GRAS1 enrichment measured by RT-qPCR after immunoprecipitation on STAT5 in A549 cell lysis. Error bars represent SD. (*) $P \leq 0.05$; (***) $P < 0.001$ by two-tailed Student's t-test.

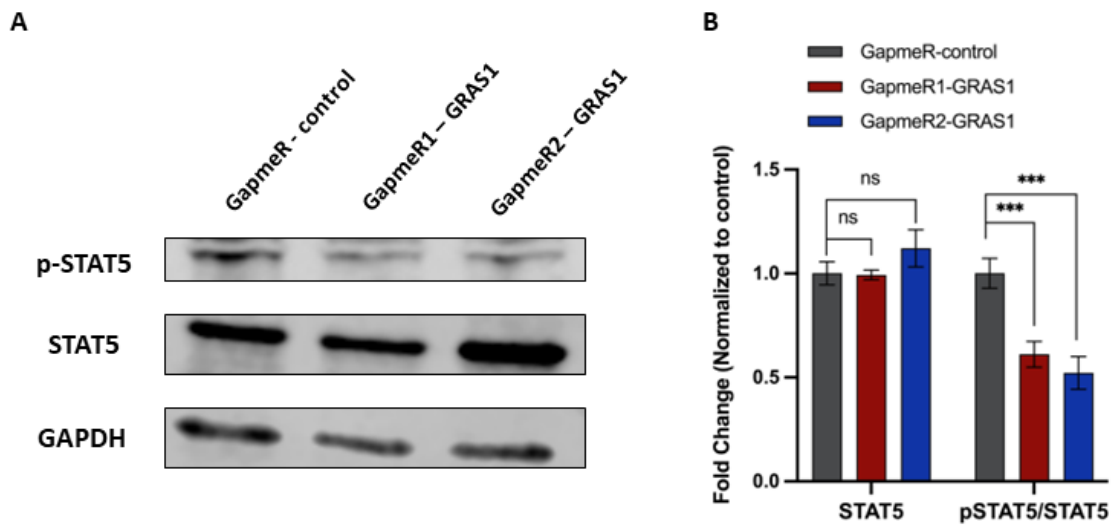


Figure 3.5 Knockdown of GRAS1 affects phosphorylation of STAT5. (A) Immunoblotting probed for the protein level of phosphorylated STAT5 (p-STAT5) and STAT5 in K562 comparing control with GRAS1 knockdown (left). Quantification of the protein level of STAT5 and p-STAT5 by normalizing to GAPDH (right). Error bars represent SD. (ns) $P > 0.05$; (***) $P < 0.001$ by two-tailed Student's t-test.

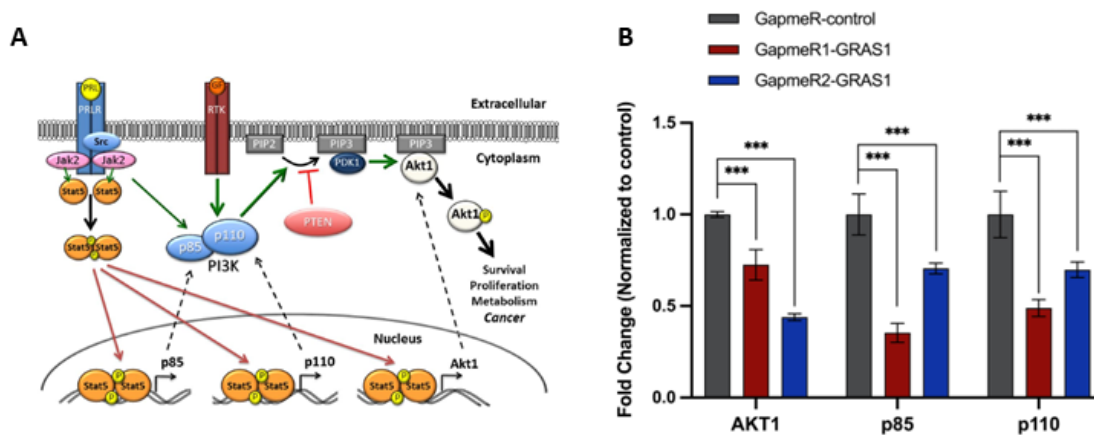


Figure 3.6 Knockdown of GRAS1 affects downstream gene targets of STAT5 signaling pathway. (A) STAT5 signaling pathway reported by a study (Schmidt et al. 2014). (B) mRNA expression of AKT1, p85, and p110 in A549 cells comparing control with GRAS1 knockdown. Error bars represent SD. (***) $P < 0.001$ by two-tailed Student's t-test.

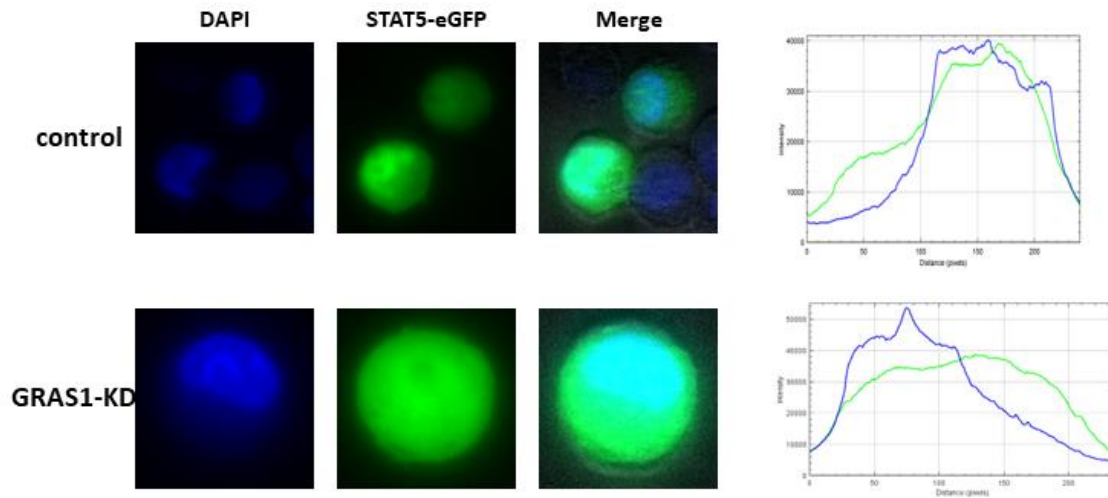


Figure 3.7 Knockdown of GRAS1 affects subcellular localization of STAT5. Both control and GRAS1 knockdown samples were transfected with eGFP-tagged STAT5 vector, followed by imaging under a confocal microscope. DAPI staining (blue) and eGFP signaling (green) are in the images. Shown are representative images (left) and quantification analysis in the experiment (right).

3.4 Discussion/future directions

In this study, we found GRAS1 directly interacts with STAT5 and regulates STAT5 phosphorylation activity in K562 cells. Based on the result of MTT assay screening on GRAS1, we found that altered expression of GRAS1 led to significant effects on cell viability in A549. Using different methods for measuring cell survival, including MTT assay, CCK6 cell survival detection kit, and annexin V apoptotic assay, we found GRAS1 does not affect cell survival in K562 cells.

A number of studies have shown lncRNAs can act as scaffolds to recruit functional proteins and subsequently affect biological processes (Roberts et al. 2014). Identification of GRAS1 protein interacting partners was required to gain a mechanistic insight on how GRAS1 regulates cell survival in A549. In general, RNA-binding proteins are identified primarily through RNA pulldown followed by mass spectrometry (Graindorge et al. 2019). RAP-MS is highly selective tool for identifying specific and direct protein interacting partners with RNA molecules, which was previously used successfully to identify the XIST non-coding RNA interaction with SPEN/SHARP in initiation of X-chromosome silencing (McHugh et al. 2015). Despite the big advantages in RAP-MS technique, a large amount of cells are required as input for this technique on lncRNA, because most lncRNAs have a low cellular abundance (Tsagakis et al. 2020). A549 is a type of adherent cell, which would be time-consuming to collect millions of cells for the RAP-MS. K562 is a type of suspension cell that could be easily grown in a big bucket. Since proteins that directly bind to U1 are known from other studies (So et al, 2019), we used RAP-MS to pull down U1 and identify U1 interacting proteins in different cell lines, including K562, A549, and HCT116. Our result of the RAP-MS indicated top U1 interacting proteins based on the MS intensity were largely nuclear ribonucleoproteins among these cell

lines. Besides, U1 binds to the nuclear ribonucleoproteins to functions in splicing introns and telescripting (Oh et al. 2017), so these U1 experiments suggested RAP-MS is sufficient to identify important protein interacting factors based on intensity signals from the mass spectrometry without the use of SILAC or other specific protein labeling methods. We did the RAP-MS on GRAS1 two times in K562, we found STAT5 was among the proteins with highest MS intensity in both RAP-MS data. The interaction between STAT5 and GRAS1 was also validated by RNA immunoprecipitation. Surprisingly, STAT5 has biological functions in the immune system (Verdeil et al. 2019, Rani and Murphy. 2016), leukemia (Shuai et al. 1996, Warsch et al. 2011), and breast cancer (Walker et al, 2009), however, few studies suggest that STAT5 is related to lung cancer. U1 is a conserved non-coding RNA that has similar functions in different cell lines, so it is not surprising to see its protein interacting partners are similar. However, GRAS1 functions differently between A549 and K562, so the assumption that the lncRNA share the similar interactomes in these cell lines was invalid. A further study may be needed to test a hypothesis that whether functional specificity in lncRNAs is caused by distinct RNA-protein interactions in different cell lines.

Although STAT5 is not the protein target to explain how GRAS1 regulates cell viability in A549, the finding that STAT5 is an RNA-binding protein is intriguing. STAT5 is a well-known transcriptional factor that is able to bind DNA for regulation of specific gene expressions (Calò et al. 2003), but its interaction with RNA is poorly addressed. In addition, we found downregulation of GRAS1 by GapmeR knockdown affects phosphorylation level of STAT5 without affecting the total level of STAT5 in K562 cells. Multiple studies have demonstrated that phosphorylation plays an important role in STAT5 translocation from cytoplasm to nucleus

(Schmidt et al, 2014). Consistently, we also found that downregulation of GRAS1 affects subcellular localization of STAT5 proteins in K562 cells.

Here, we reported that GRAS1 does not affect cell viability in K562 cells, but we found that knockdown of GRAS1 using either of the GapmeRs changed culture medium color from red to yellow because of alteration of acidification in the medium, suggesting that GRAS1 affects cellular metabolism in K562 by increasing extracellular lactate production. Quantitative biochemical assays are needed to measure metabolic alteration. It may provide a potential connection with the biological significance of the GRAS1-STAT5 nucleoprotein complex.

3.5 Reference:

1. Yoon, J. H., Abdelmohsen, K., & Gorospe, M. (2013). Posttranscriptional gene regulation by long noncoding RNA. *Journal of molecular biology*, 425(19), 3723-3730.
2. Bhan, A., & Mandal, S. S. (2014). Long noncoding RNAs: emerging stars in gene regulation, epigenetics and human disease. *ChemMedChem*, 9(9), 1932-1956.
3. Huarte, M. (2015). The emerging role of lncRNAs in cancer. *Nature medicine*, 21(11), 1253-1261.
4. He, R. Z., Luo, D. X., & Mo, Y. Y. (2019). Emerging roles of lncRNAs in the post-transcriptional regulation in cancer. *Genes & diseases*, 6(1), 6-15.
5. Chen, X., Sun, Y., Cai, R., Wang, G., Shu, X., & Pang, W. (2018). Long noncoding RNA: multiple players in gene expression. *BMB reports*, 51(6), 280.
6. Holoch, D., & Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. *Nature Reviews Genetics*, 16(2), 71-84.

7. Koziol, M. J., & Rinn, J. L. (2010). RNA traffic control of chromatin complexes. *Current opinion in genetics & development*, 20(2), 142-148.
8. Wang, M., Zheng, S., Li, X., Ding, Y., Zhang, M., Lin, L., Xu, H., Cheng, Y., Zhang, X., Xu, H. and Li, S., 2020. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network identified lncRNA EPB41L4A-AS1 as a potential biomarker in non-small cell lung cancer. *Frontiers in Genetics*, 11, p.511676.
9. Delshad, E., Shamsabadi, F. T., Bahramian, S., Mehravar, F., Maghsoudi, H., & Shafiee, M. (2019). In silico identification of novel lncRNAs with a potential role in diagnosis of gastric cancer. *Journal of Biomolecular Structure and Dynamics*.
10. Bin, J., Nie, S., Tang, Z., Kang, A., Fu, Z., Hu, Y., Liao, Q., Xiong, W., Zhou, Y., Tang, Y. and Jiang, J., 2021. Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating the Rho/ROCK pathway in colorectal cancer. *Journal of cellular physiology*, 236(1), pp.523-535.
11. Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., Wang, Y., Wang, S., Zhu, Y., Chen, D. and Xie, W., 2019. LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. *EBioMedicine*, 41, pp.200-213.
12. Zhu, Y., Liu, Q., Liao, M., Diao, L., Wu, T., Liao, W., Wang, Z., Li, B., Zhang, S., Wang, S. and Xie, W., 2019. Overexpression of lncRNA EPB41L4A-AS1 induces metabolic reprogramming in trophoblast cells and placenta tissue of miscarriage. *Molecular Therapy-Nucleic Acids*, 18, pp.518-532.
13. Cui, P., Zhao, X., Liu, J., Chen, X., Gao, Y., Tao, K., Wang, C. and Zhang, X., 2020. miR-146a interacting with lncRNA EPB41L4A-AS1 and lncRNA SNHG7 inhibits proliferation of

- bone marrow-derived mesenchymal stem cells. *Journal of Cellular Physiology*, 235(4), pp.3292-3308.
14. Roberts, T. C., Morris, K. V., & Weinberg, M. S. (2014). Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*, 9(1), 13-20.
 15. McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A. and Sweredoski, M.J., 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521(7551), pp.232-236.
 16. McHugh, C. A., Russell, P., & Guttman, M. (2014). Methods for comprehensive experimental identification of RNA-protein interactions. *Genome biology*, 15, 1-10.
 17. Gottlieb E, Rinn JL (2016) RNA-protein interactions in gene regulation. *Nature Reviews Genetics* 17: 703-716.
 18. Calò, V., Migliavacca, M., Bazan, V., Macaluso, M., Buscemi, M., Gebbia, N., & Russo, A. (2003). STAT proteins: from normal control of cellular events to tumorigenesis. *Journal of cellular physiology*, 197(2), 157-168.
 19. Coffey, P. J., Koenderman, L., & de Groot, R. P. (2000). The role of STATs in myeloid differentiation and leukemia. *Oncogene*, 19(21), 2511-2522.
 20. de Groot, R. P., Coffey, P. J., & Koenderman, L. (1998). Regulation of proliferation, differentiation and survival by the IL-3/IL-5/GM-CSF receptor family. *Cellular signalling*, 10(9), 619-628.
 21. Buettner, R., Mora, L. B., & Jove, R. (2002). Activated STAT signaling in human tumors provides novel molecular targets for therapeutic intervention. *Clinical cancer research*, 8(4), 945-954.

22. Halim, C. E., Deng, S., Ong, M. S., & Yap, C. T. (2020). Involvement of STAT5 in Oncogenesis. *Biomedicines*, 8(9), 316.
23. Rane, S. G., & Reddy, E. P. (2000). Janus kinases: components of multiple signaling pathways. *Oncogene*, 19(49), 5662-5679.
24. Cella, N., Groner, B., & Hynes, N. E. (1998). Characterization of Stat5a and Stat5b homodimers and heterodimers and their association with the glucocorticoid receptor in mammary cells. *Molecular and Cellular Biology*, 18(4), 1783-1792.
25. Paukku, K., & Silvennoinen, O. (2004). STATs as critical mediators of signal transduction and transcription: lessons learned from STAT5. *Cytokine & growth factor reviews*, 15(6), 435-455.
26. So, Byung Ran, Chao Di, Zhiqiang Cai, Christopher C. Venters, Jiannan Guo, Jung-Min Oh, Chie Arai, and Gideon Dreyfuss. "A complex of U1 snRNP with cleavage and polyadenylation factors controls telescripting, regulating mRNA transcription in human cells." *Molecular cell* 76, no. 4 (2019): 590-599.
27. Harir, N., Boudot, C., Friedbichler, K., Sonneck, K., Kondo, R., Martin-Lannerée, S., Kenner, L., Kerenyi, M., Yahiaoui, S., Gouilleux-Gruart, V. and Gondry, J., 2008. Oncogenic Kit controls neoplastic mast cell growth through a Stat5/PI3-kinase signaling cascade. *Blood, The Journal of the American Society of Hematology*, 112(6), pp.2463-2473.
28. Matsumura, Itaru, Toshio Kitamura, Hiroshi Wakao, Hirokazu Tanaka, Koji Hashimoto, Chris Albanese, Julian Downward, Richard G. Pestell, and Yuzuru Kanakura. "Transcriptional regulation of the cyclin D1 promoter by STAT5: its involvement in cytokine-dependent growth of hematopoietic cells." *The EMBO journal* 18, no. 5 (1999): 1367-1377.

29. Verdeil, Grégory, Toby Lawrence, Anne-Marie Schmitt-Verhulst, and Nathalie Auphan-Anezin. "Targeting STAT3 and STAT5 in tumor-associated immune cells to improve immunotherapy." *Cancers* 11, no. 12 (2019): 1832.
30. Rani, Aradhana, and John J. Murphy. "STAT5 in cancer and immunity." *Journal of Interferon & Cytokine Research* 36, no. 4 (2016): 226-237.
31. Shuai, Ke, Jocelyn Halpern, J. Ten Hoeve, Xiaoping Rao, and Charles L. Sawyers. "Constitutive activation of STAT5 by the BCR-ABL oncogene in chronic myelogenous leukemia." *Oncogene* 13, no. 2 (1996): 247-254.
32. Warsch, W., Kollmann, K., Eckelhart, E., Fajmann, S., Cerny-Reiterer, S., Hölbl, A., Gleixner, K.V., Dworzak, M., Mayerhofer, M., Hoermann, G. and Herrmann, H., 2011. High STAT5 levels mediate imatinib resistance and indicate disease progression in chronic myeloid leukemia. *Blood, The Journal of the American Society of Hematology*, 117(12), pp.3409-3420.
33. Walker, Sarah R., Erik A. Nelson, Lihua Zou, Mousumi Chaudhury, Sabina Signoretti, Andrea Richardson, and David A. Frank. "Reciprocal effects of STAT5 and STAT3 in breast cancer." *Molecular Cancer Research* 7, no. 6 (2009): 966-976.
34. Stoecklin, Elisabeth, Manuela Wissler, Richard Moriggl, and Bernd Groner. "Specific DNA binding of Stat5, but not of glucocorticoid receptor, is required for their functional cooperation in the regulation of gene transcription." *Molecular and Cellular Biology* 17, no. 11 (1997): 6708-6716.
35. Schmidt, J.W., Wehde, B.L., Sakamoto, K., Triplett, A.A., Anderson, S.M., Tschlis, P.N., Leone, G. and Wagner, K.U., 2014. Stat5 regulates the phosphatidylinositol 3-kinase/Akt1

pathway during mammary gland development and tumorigenesis. *Molecular and cellular biology*, 34(7), pp.1363-1377.

36. Schmidt, Jeffrey W., Barbara L. Wehde, Kazuhito Sakamoto, Aleata A. Triplett, Steven M. Anderson, Philip N. Tschlis, Gustavo Leone, and Kay-Uwe Wagner. "Stat5 regulates the phosphatidylinositol 3-kinase/Akt1 pathway during mammary gland development and tumorigenesis." *Molecular and cellular biology* 34, no. 7 (2014): 1363-1377.
37. Graindorge, A., Pinheiro, I., Nawrocka, A., Mallory, A.C., Tsvetkov, P., Gil, N., Carolis, C., Buchholz, F., Ulitsky, I., Heard, E. and Taipale, M., 2019. In-cell identification and measurement of RNA-protein interactions. *Nature communications*, 10(1), p.5317.
38. Oh, J.M., Di, C., Venters, C.C., Guo, J., Arai, C., So, B.R., Pinto, A.M., Zhang, Z., Wan, L., Younis, I. and Dreyfuss, G., 2017. U1 snRNP telescripting regulates a size–function-stratified human genome. *Nature structural & molecular biology*, 24(11), pp.993-999.
39. Koch, L. (2017). Screening for lncRNA function. *Nature Reviews Genetics*, 18(2), 70-70.
40. Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y. and Mandegar, M.A., 2017. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, 355(6320), p.eaah7111.

3.6 Acknowledgement

Chapter 3, in full, is currently being prepared for publication of the material. The authors in the publication: Tong Su, Lingbo Kong, Darin Cheung, Calvin Huang, and Colleen McHugh. The dissertation author is the primary investigator and author of this material.

CHAPTER 4: GRAS1 long non-coding RNA protects cells from DNA damage and ferroptosis

4.1 Abstract

Non-coding RNAs have been implicated as regulators of cell growth and survival in a variety of cancer types, but the specific biological functions and interaction partners of many non-coding RNAs are still not well understood. We characterized the function of the non-coding RNA Growth Regulator Antisense 1 (GRAS1) in non-small cell lung cancer cells. GRAS1 is critical for A549 cell survival, and depletion of GRAS1 RNA transcript induces cell death. We found that knockdown of GRAS1 resulted in a significant increase in protein levels of p53 and the DNA damage marker γ H2AX. GRAS1 knockdown leads to differential expression of many mRNA transcript levels, including upregulation of DNA damage response factors involved in double strand break repair. Further, we identified the NF- κ B activating protein (NKAP) as a direct interaction partner of GRAS1 using RNA antisense purification with mass spectrometry (RAP-MS). Decreased GRAS1 levels in lung cancer cells caused a significant decline in NKAP level and induction of cell death through ferroptosis. The addition of the ferroptosis inhibitors ferrostatin and α -tocopherol could partially rescue the effect of GRAS1 knockdown on lung cancer cell survival. We found NKAP overexpression could completely reverse the ferroptosis as well as rescue the cell viability promoted by GRAS1 downregulation in lung cancer cells. Besides, NKAP overexpression could significantly reverse the effect of GRAS1 knockdown on lung cancer DNA damage.

4.2 Introduction

Lung cancer is one of the leading causes of cancer-related deaths worldwide (Veena et al. 2013). It is a tumor that grows in the lungs and can spread to other parts of the body in the late stage. Most lung cancer cases are caused by long-term exposure to tobacco smoke (de Groot et al. 2018), but other factors such as air pollution, occupational exposure to carcinogens, and genetic predisposition can also increase the risk of developing the disease. According to a report from the World Health Organization (2020), lung cancer is the most diagnosed cancer worldwide, with an estimated 2.2 million new cases and 1.8 million deaths in 2020 alone. Lung cancers generally fall under two main types: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC accounts for about 85% of all lung cancer cases, while SCLC accounts for the remaining 15% (Mayo Clinic, 2021). The survival rate for lung cancer is generally poor, with only about 19% of patients surviving for five years or more following diagnosis (National Cancer Institution, 2021). In addition, between 30% to 55% of patients with NSCLC experience a recurrence (Uramoto and Tanaka. 2014). Current treatment options for lung cancer include surgery, radiation therapy, chemotherapy, targeted therapy, and immunotherapy (Mayo Clinic, 2021).

LncRNAs are emerging as important regulators of gene expression and cellular processes, including tumorigenesis (Kazemzadeh et al. 2015). Recently, numerous studies have shown that lncRNAs can act as oncogenes, promoting cell proliferation and survival, or as tumor suppressors, suppressing cell growth and division (Wang and Chang, 2011). Further research into the role of lncRNAs in cancer has the potential to discover new therapeutic targets and treatments for this disease. Understanding the mechanisms by which lncRNAs contribute to

cancer will provide new insights into the underlying biology of this disease and could lead to the development of more effective and personalized therapies.

GRAS1 is a long non-coding RNA that has been implicated in various types of cancer (Wang et al. 2020). It has been shown to be dysregulated in several types of cancer, including gastric cancer (Delshad et al. 2019), lung cancer (Wang et al. 2020), and colorectal cancer (Bin et al. 2021), and is associated with tumor progression and poor prognosis (Bin et al. 2021). Regarding the functional role of GRAS1, two studies have reported that GRAS1 is involved in the regulation of cellular glycolysis and glutaminolysis through inhibiting the Warburg effect in cancer cells (Liao et al. 2019) and placental villus (Zhu et al. 2019). In these studies, they found that knockdown of GRAS1 disrupts the interaction between HDAC2 and NPM1, such that HDAC2 can be released from nucleolus and bind to VHL and VDAC1 promoter region to promote cellular level of glycolysis and glutaminolysis (Liao et al. 2019). A recent study suggests GRAS1 regulates cellular proliferation and cell cycle phase distribution in bone marrow-derived mesenchymal stem cells (Cui et al. 2020). Taken together, GRAS1 is suggested to play an important role in regulating cell growth in cancer. In this study, we investigated the role of GRAS1 in A549 lung cancer cells.

A number of studies have demonstrated that the NF- κ B signaling pathway plays a crucial role in the regulation of various physiological processes, including inflammation (Sun 2017), immune responses (Hayden et al. 2006), and cell survival (Piva et al. 2006). When the NF- κ B pathway is dysregulated, it can contribute to the development and progression of cancer (Matthews et al. 2016). NF- κ B activating protein (NKAP) is a nuclear localized protein that acts as a regulator of tumor necrosis factor- and interleukin-1-induced NF- κ B activation (Chen et al. 2003). NKAP was first identified by a yeast two-hybrid screening for potential receptor-

interacting protein (RIP) interacting proteins in 2003 (Chen et al. 2003). NKAP is a multifunctional protein and is essential for growth and development processes in multiple highly proliferative cell types including spermatozoa, T cells, and hematopoietic stem cells. (Pajerowski et al. 2009, Shapiro et al. 2019, Okuda et al. 2015). NKAP has also been shown to bind to both mRNA and pre-mRNA, and it plays an important role in RNA splicing and processing (Burgute et al. 2014). Furthermore, NKAP functions as a cancer-promoting oncogene and has shown significant phenotypes in many types of cancer, including gastric cancer (Wei et al. 2019), colon cancer (Shu et al. 2019), renal cell carcinoma (Ma et al. 2020), but its role in lung cancer remains largely unidentified. In addition, NKAP is shown to act as a key regulator of mitosis via its activity of SUMOylation, and its dysregulation can cause chromosomal instability in cancer cells (Li et al. 2016). Lastly, recent research reported that knockdown of NKAP can induce ferroptosis in glioblastoma cell lines and has been shown to protect cells from ferroptosis by acting as a new ferroptosis suppressor (Sun et al. 2022).

Since 2012, ferroptosis has emerged as a relatively newly discovered iron-dependent, non-apoptotic pathway for cell death (Dixon and Stockwell. 2019). Ferroptosis has been uniquely characterized as morphologically and biochemically distinct from apoptosis, necrosis, and autophagy (Dixon and Stockwell. 2019), and regulated by signaling pathways such as p53, Nrf2, and FANCD2 (Xu et al. 2019). Ferroptosis has unique features that distinguish ferroptosis from other types of cell death, including lipid peroxidation, large free iron buildup, and mitochondrial defects such as impacted cristae and outer membrane ruptures (Li et al. 2020). During ferroptosis, an imbalance of intracellular iron leads to accumulation of reactive oxygen species (ROS), which leads to abnormally high concentrations of lipid peroxidase generation, ultimately leading to loss of membrane permeability and cell death (Xu et al. 2019). Ferroptosis

is characterized by two specific hallmarks: an external regulation of iron concentration via the transferrin receptor pathway, and an internal enzymatic redox pathway that involves glutathione and GPX4 in reducing lipid peroxidases (Chen et al. 2021). Each process is critical in controlling the levels of lipid peroxidases in order to prevent internal cellular damage from highly reactive peroxidase reactions. Typically, iron concentrations are regulated by binding to PCBP1 and stored as ferritin in the lysosome, and exported by ferroportin transporters, which prevent free iron from reacting to other biochemical molecules (Xu et al. 2019). One such reaction is the conversion of polyunsaturated fatty acids (PUFA) into polyunsaturated acid peroxidases (PUFA-PE), which requires iron as a cofactor (Xu et al. 2019). The accumulation of PUFA-PE is harmful to the cell, and the buildup of lipid peroxidases will ultimately result in oxidative lipid membrane degradation (Chen et al. 2021).

In this study, we found that knockdown of GRAS1 by GapmeRs can significantly decrease the cell survival in lung cancer cells while generating a large degree of DNA damage. Furthermore, the downregulation of GRAS1 can cause dramatic activation of the p53 pathway and disruption of cell cycle phase distribution. In the context of its biological mechanism in lung cancer, we found that GRAS1 is directly interacting with NKAP protein to protect the cell from ferroptosis in lung cancer cells. When GRAS1 is downregulated, protein levels of the NKAP decrease dramatically, which leads to a decrease in the NF- κ B signaling pathway as well as an increase in cellular ferroptosis. We found NKAP overexpression could completely reverse the ferroptosis as well as rescue the cell viability promoted by GRAS1 downregulation in lung cancer cells. Besides, NKAP overexpression could significantly reverse the effect of GRAS1 knockdown on lung cancer DNA damage.

4.3 Results

4.3.1 GRAS1 is required for cell survival in A549

GRAS1 is a lncRNA that is in chromosome 5 region with 1194 nucleotides in length, however, only part of GRAS1 is expressed in cells based on the layered RNA sequencing analysis across different cell lines from UCSC Encode Project (Figure 4.1A). GRAS1 is well expressed in many different tissues (Figure 4.1B). To address the question of GRAS1 subcellular localization in lung cancer cells, WT A549 cells were separated into cytoplasmic and nuclear fractions via Gagnon fractionation, and RNA was extracted from each fraction for analysis by RT-qPCR. GAPDH and U1 were amplified as cytoplasmic and nuclear standards respectively. As shown in Figure 4.1C, GRAS1 was detected in both fractions with similar distribution, while the controls were localized as expected.

Since GRAS1 is dysregulated in different types of cancer based on multiple studies, we intended to determine whether GRAS1 regulates cell viability in lung cancer cells. To do that, we generated a system for overexpression and knockdown of GRAS1 in A549. GRAS1 gene was cloned and inserted into pcDNA by the Gateway Cloning System. As shown in Figure 4.2A, using the pcDNA-GRAS1 plasmid could efficiently overexpress GRAS1 compared with the control vector in A549. For the knockdown of GRAS1, we designed two GapmeRs to target the GRAS1, referred as GapmeR1 and GapmeR2, to eliminate any possible off-target effects that may cause by the GapmeR itself (estimate target sites of both GapmeRs on GRAS1 are shown at Figure 4.2B). Using RT-qPCR, we found an average GRAS1 knockdown of ~60% was achieved in A549 cells transfected with either GapmeR1 or GapmeR2 compared with control GapmeR (Figure 4.2B).

Based on the MTT assay screening using pcDNA3.2-GRAS1 vector in different cell lines, we found that overexpression of GRAS1 could significantly increase the cell survival rate of A549 cells (Figure 4.3A, 4.3C). To further validate this result, colony formation assay was used as an alternative method to observe A549 survival rate. As shown in Figure 4.3D, ectopic expression of GRAS1 in A549 cells enhanced colony formation compared to the vector group. On the contrary, we found that knockdown of GRAS1 could dramatically decrease the cell survival rate of A549 using GapmeR1 (Figure 4.3A), based on our MTT assay screening in various cell lines. To eliminate any off-target effects caused by GapmeR1, we repeated MTT assay on A549 cells that were treated with either GapmeR1 or GapmeR2 and compared the cell viability with the control GapmeR. Based on the MTT assay, knockdown of GRAS1 using either GapmeR1 or GapmeR2 could significantly decrease the cell survival to a similar degree in A549 (Figure 4.3B). The colony formation assay was repeated to validate the result of the MTT assay. As shown in Figure 4.3E, knockdown of GRAS1 in A549 cells dramatically decreases colony formation compared to the control group. In conclusion, GRAS1 is critical for A549 cell survival, and depletion of GRAS1 RNA transcript induces cell death.

4.3.2 RNA sequencing identifies transcriptomic programs affected by GRAS1 knockdown

Since knockdown of GRAS1 led to a dramatic decrease in cell survival, we were interested in determining whether any transcriptional program alterations occurs when GRAS1 is downregulated by GapmeR knockdown in A549. RNA sequencing was performed on total RNA extracted from GRAS1 knockdown cells. Principal Component Analysis (PCA) was performed to validate the quality of the RNA sequencing and to visualize the differences between groups as well as samples in the same group. As shown in Figure 4.4A, the two groups, GRAS1

knockdown and GapmeR control, showed distinct differences between each other, while each group of RNA samples was consistent among the biological triplicates. Furthermore, knockdown efficiency of GRAS1 was confirmed using normalized read counts in the RNAseq data (Figure 4.4B). We found that EPB41L4A, a GRAS1 nearby gene, and SNORA13, a snoRNA generated from GRAS1 intron, had no significant changes in their mRNA level when GRAS1 was downregulated (Figure 4.4B).

Based on the overview of the RNAseq data in volcano plotting, downregulation of GRAS1 by GapmeR knockdown led to significant expression alterations in many different genes (Figure 4.4C). To further analyze the RNAseq data, we identified a total of 12062 genes that are more than 2-fold differentially expressed with $p\text{-value} \leq 0.05$, including 41.4% upregulation and 58.6% downregulation overall. Since knockdown of GRAS1 induced too many differentially expressed genes in A549, we selected the top 1000 genes based on their p-value to maximize confidence for RNAseq data analysis. Gene ontology (GO) enrichment analysis of these differentially expressed genes exhibited enrichment of various signaling pathways linked with cell survival, including p53 signaling pathway, DNA damage, and cell cycle phase distribution (Figure 4.4D). To further assess these cell-survival related pathways, heatmaps of the top 20 genes sorted based on the p-value in the three pathways were shown in Figure 4.4E. Selected genes in all those pathways were significantly upregulated. In conclusion, GRAS1 knockdown leads to differential expression of many mRNA transcript levels, including multiple cell-survival related pathways, including p53 signaling pathway, DNA damage, and cell cycle phase distribution.

4.3.3 Knockdown of GRAS1 induces p53 activation in A549

P53 is a key regulator for both cell viability and cell cycle phase distribution (Kastan et al. 1995). Since a strong upregulation of p53 was observed on the heatmap analysis, we investigated whether GRAS1 regulates the p53 signaling pathway to impact cell survival of A549 cells. To approach this question, we first assessed whether GRAS1 has an effect on p53 protein levels in A549 cells. As shown in Figure 4.5A, knockdown of GRAS1 using either of the GapmeRs significantly increased the protein level of p53. It has been reported that p53 turnover is regulated by ubiquitination (Brooks and Gu. 2006), so we determined whether GRAS1 is involved in p53 ubiquitination by using an in-vitro ubiquitination assay. A549 cells were transfected with either GapmeR1 or GapmeR2 against GRAS1 along with the control for 48hrs, followed by p53 pulldown using anti-p53 antibody. Ubiquitination level on p53 was detected by the anti-ubiquitin antibody. As shown in Figure 4.5B, knockdown of GRAS1 had no effect on p53 ubiquitination level. To further assess the effect of GRAS1 knockdown on the p53 pathway, typical p53 downstream gene targets, including p21, PML, PLK3, and Gadd45a expression were examined by qPCR at A549. Consistently, knockdown of GRAS1 using either of the GapmeRs dramatically increased mRNA expression of all p53 downstream targets (Figure 4.5C). In contrast, we found that overexpression of GRAS1 had no significant effects on the protein level of p53 in A549, as shown in Figure 4.6A. Similarly, there were no significant differences observed in the same p53 downstream targets as we addressed when GRAS1 got overexpressed at A549 (Figure 4.6B). These data indicated that downregulation of GRAS1 induces p53 level and activates p53 signaling pathway in lung cancer cells.

4.3.4 GRAS1 affects cell cycle distribution in A549

The cell cycle plays a critical role in maintaining cell viability by ensuring that the cell replicates and divides correctly (Matthews et al. 2022). To assess the role of GRAS1 in cell cycle regulation, we performed propidium iodide (PI) staining with flow cytometry analysis. We found downregulation of GRAS1 affects cell cycle distribution in A549 (Figure 4.7A), which is consistent with the heatmap analysis that was shown in Figure 4.4E. As shown in Figure 4.7A, downregulation of GRAS1 led to a significant increase in sub-G1 phase in A549, suggesting that the cells were going for an apoptotic state. In addition, we found that lack of GRAS1 increased the S phase as well as decreased the G1 phase in the cell cycle (Figure 4.7A). In Contrast, ectopic expression of GRAS1 led to a decrease in sub-G1 phase in A549, while no significant effects of other phases in the cell cycle were observed (Figure 4.7B). These data suggested that GRAS1 affects cell cycle distribution in A549 cells.

4.3.5 Knockdown of GRAS1 induces DNA damage in A549

DNA damage is one of the biological factors that trigger activation of the p53 signaling pathway (Lakin and Jackson, 1999). Based on the observation of a strong upregulation of the DNA damage repair pathway according to the heatmap analysis (Figure 4.4E), we investigated whether the knockdown of GRAS1 increases the cellular level of DNA damage in A549 cells. To measure the cellular level of DNA damage, immunoblotting was used to detect γ H2AX expression. γ H2AX is the histone variant H2AX with the phosphorylated Ser-139 that is caused by an early cellular response from DNA damage (Mah et al. 2010). As shown in Figure 4.8A, knockdown of GRAS1 using either of the GapmeRs significantly increased the protein expression of γ H2AX, while no significant alteration in γ H2AX level was observed when GRAS1 was upregulated in A549 cells (Figure 4.6A). To obtain a better visualization of DNA

damage intensity in the cells, we performed immunofluorescence to count γ H2AX foci using a confocal microscope. A higher number of cellular γ H2AX foci corresponds to a higher level of DNA damage in the cell. Consistent with the result of γ H2AX immunoblotting, we found that knockdown of GRAS1 using either of the GapmeRs dramatically increased the average number of γ H2AX foci compared with the control GapmeR in A549 cells (Figure 4.8B), suggesting that downregulation of GRAS1 leads to a significant increase in the level of DNA damage in A549. Chromosome condensation along with DNA fragmentation is one of the typical characteristics in apoptotic cells that is caused by DNA damage (Toné et al., 2007). Since the cellular level of DNA damage was significantly induced when GRAS1 was downregulated, we discovered that the knockdown of GRAS1 could significantly upregulate the level of chromosome condensation in A549 under the cell microscope using Hoechst staining assay (Figure 4.8C). Furthermore, in agreement with the DNA staining assay, we found that the downregulation of GRAS1 strongly enhanced the protein level of the apoptosis marker cleaved PARP. These data indicate that the downregulation of GRAS1 induces DNA damage in A549 cells.

4.3.6 GRAS1 directly interacts with NKAP

Many studies have shown that lncRNAs can act as scaffolds to recruit functional proteins and subsequently affect biological processes (Roberts et al. 2014). To investigate the biological mechanism in which GRAS1 regulates cell proliferation in A549 cells, we decided to determine the protein partners that are directly interacting with GRAS1 endogenously in A549 cells via RNA antisense purification with mass spectrometry (RAP-MS). Before performing RAP-MS, validation of designed biotinylated DNA probes for GRAS1 was performed by capturing GRAS1

RNA in the total RNA pool from A549 cell lysates and quantifying efficiency through the RAP-qPCR technique (Figure 4.9A). Since GRAS1 is not as abundant as U1 RNA in A549 cells, we increased the A549 cell input of the RAP-MS experiment on GRAS1 to enrich the protein capture for mass spectrometry. Thus, 150M A549 cells for GRAS1, along with 50M A549 cells for luciferase as a negative control and 50M A549 cells for U1 as a positive control, were used for the pull-down. 53 proteins, excluding the same ones from the luciferase group as non-specific interacting partners, with a minimum of 2 unique peptides were pulled down and identified in the GRAS1 group. The majority of these interactomes were cytoplasmic proteins, while 27 of them were in both cytoplasmic fraction and nuclear fraction (Figure 4.9B). Since GRAS1 functions differently compared with U1 in cells, only 18 proteins were found that were specifically GRAS1 binding proteins when the common proteins with U1 were filtered (Figure 4.9C). Among these 18 proteins, we found multiple mitochondrial proteins involved in the cellular metabolism (GANAB, TUFM, P4HB, PDIA4, MDH2, PCCB, ATP5B), a few heat shock proteins (HSPE1, HSPA4, HSP90) and one protein involved in cell growth signaling pathway (NKAP).

Since A549 cells lacking GRAS1 undergo apoptosis from DNA damage and p53 activation, we focused on NKAP, which is shown to be involved in the cellular regulation of ferroptosis (Sun et al. 2022) and centrosomes (Li et al. 2016) that could link to DNA instability. Interestingly, NKAP and NKAP domain containing 1 (NKAPD1) were both found as the top proteins based on MS intensity in the list of RAP-MS data. Despite NKAP and NKAPD1 having similar names, these two genes have distinct biological functions and are not similar in terms of protein sequence or structure. Using the NCBI Protein Blast searching tool, there was no significant similarity between NKAP and NKAPD1 based on their protein sequences. In terms of

biological function, NKAPD1 is still largely uncharacterized. Thus, we decided to focus on the NKAP protein and investigate whether the direct interaction between NKAP and GRAS1 is related to the effects of cell survival in A549 cells caused by the knockdown of GRAS1.

RNA immunoprecipitation (RIP) was used to confirm the interaction between GRAS1 and NKAP. Using a commercial antibody against NKAP in 30M A549 protein lysis to pull down NKAP protein along with an antibody against IgG as a negative control followed by qPCR, we found over a 50-fold increase on GRAS1 capture in NKAP-IP elution compared with the IgG-IP group (Figure 4.9D). These data from both RAP-MS and RIP indicated that GRAS1 is binding with NKAP in A549.

4.3.7 Knockdown of GRAS1 affects cellular activity of NF- κ B pathway

Since GRAS1 interacts with NKAP in A549, we investigated the biological role of this interaction. Surprisingly, we found that the knockdown of GRAS1 led to a significant decrease in the protein level of the NKAP (Figure 4.11A). Using RT-qPCR, we found that the knockdown of GRAS1 had no significant change in the mRNA level of NKAP (Figure 4.12A), suggesting GRAS1 may post-translationally regulate NKAP. In addition, MG132 treatment after GRAS1 got down-regulated completely rescued the decrease in the protein level of NKAP compared with the control (Figure 4.12B), suggesting GRAS1 may protect NKAP from ubiquitination.

One of the proposed functions of NKAP is that it can cooperate with cellular substrates to activate the NF- κ B pathway (Liu et al, 2021), we decided to examine whether the decrease in the protein level of NKAP caused by the lack of GRAS1 affects the cellular activity of the NF- κ B signaling pathway. Using RT-qPCR, we found that downstream mRNA targets of the NF- κ B pathway, including GCLC, ANGPT1, BDKRB1, NR3C1, and BACE1, were all down-regulated

(Figure 4.11B) when the GRAS1 got knocked down using either of the GapmeRs, suggesting knockdown of GRAS1 can decrease the cellular activity of NF- κ B pathway in A549.

Furthermore, we found a dramatic decrease in the protein level of NF- κ B protein (p65) when the GRAS1 got knocked down in A549 (Figure 4.11A). To eliminate any possible off-target effects on protein expression that are caused by the GRAS1 knockdown, we examine levels of a few nonspecific proteins, including an apoptosis-related protein (Cytochrome-C), and a critical lung cancer-associated metabolic protein (HIF1a). Based on the immunoblotting, we did not see any significant changes in either of the proteins compared with the control (only Cytochrome-C is shown in Figure 4.11A). Furthermore, we found no significant changes in the protein level of the NIK and p65, as well as the downstream targets of the NF- κ B, when the GRAS1 got overexpressed in A549. These data suggested the knockdown of GRAS1 decreases the cellular activity of the NF- κ B pathway.

Because the knockdown of GRAS1 negatively affects the cellular activity of the NF- κ B pathway, we hypothesized that activation of the NF- κ B pathway is able to rescue cell apoptosis that is caused by GRAS1 knockdown in A549. TNF- α is a cytokine that is able to activate the NF- κ B pathway (Tang et al, 2017). WT A549 cells had no significant effect on cell survival under the 24h TNF- α treatment. However, we found no significant effect in cell survival on GRAS1 knockdown in A549 with the TNF- α treatment compared to the normal condition (Figure 4.11C), suggesting that the decrease in cell viability by GRAS1 downregulation is not directly related to cellular activity of NF- κ B pathway.

4.3.8 GRAS1 affects cellular level of ferroptosis in A549

Because knockdown of GRAS1 decreases the protein level of NKAP in A549, we decided to investigate whether GRAS1 knockdown affects the cellular level of ferroptosis. Ferroptosis is another proposed function linked with NKAP (Sun et al, 2022). One of the key characteristics of ferroptosis is lipid oxidation in cell membranes, which can be detected by using BODIPY-C11 probes and flow cytometry (Martinez et al, 2020). We found that the knockdown of GRAS1 using either of the GapmeRs led to a big increase in the cellular level of ferroptosis compared with the control in A549 (Figure 4.13A). Ferrostatin and α -tocopherol are potent inhibitors for ferroptosis (Sun et al, 2022). With 24h treatment using either Ferrostatin (Figure 4.13B) or α -tocopherol (Figure 4.13C) can completely reverse the effect of ferroptosis that was caused by GRAS1 knockdown. Interestingly, the addition of the ferroptosis inhibitors ferrostatin and α -tocopherol could partially rescue the effect of GRAS1 knockdown on lung cancer cell survival (Figure 4.13E). Although the rescue effect in cell survival using either of the ferroptosis inhibitors was not 100%, it can clearly show that ferroptosis is one of the big factors in apoptosis that is caused by GRAS1 knockdown in A549. These data demonstrated that lack of GRAS1 induces ferroptosis in A549, and the addition of ferroptosis inhibitors can reverse the cellular level of ferroptosis and partially rescue the effect on cell viability caused by GRAS1 downregulation.

Surprisingly, we found a significant decrease in the cellular level of ferroptosis when the GRAS1 got overexpressed at A549 (Figure 4.13D). Overexpression of GRAS1 increased cell survival in A549, however, it had no significant effects on p53 signaling pathway and NF- κ B signaling pathway. To better understand whether ferroptosis inhibition is the factor to result in the increase in cell survival caused by ectopic expression of GRAS1, ferroptosis inducers, including Erastin and RLS, were used to activate ferroptosis. Comparing with the DMSO

control, we found that WT A549 had no change on cell survival under treatment of the Erastin, while it had a small but statistically significant decrease on cell survival under treatment of the RLS (Figure 4.13F). As shown in Figure 4.13F, overexpression of GRAS1 showed no significant effect on cell survival in A549 under treatment of either Erastin or RLS, suggesting activation of ferroptosis is able to eliminate the increase in cell survival that is caused by GRAS1 overexpression.

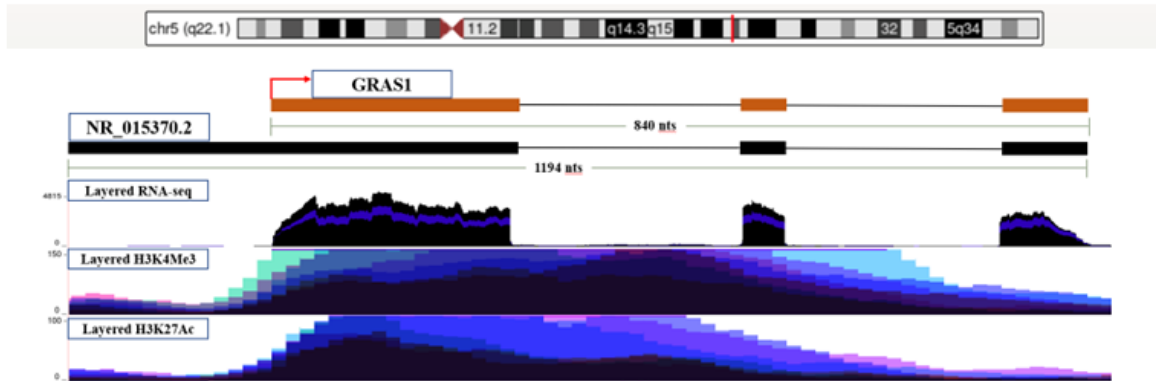
4.3.9 Overexpression of NKAP rescues A549 cells from ferroptosis and DNA damage caused by GRAS1 knockdown

Although multiple studies have demonstrated NKAP is critical for cancer development, we decided to investigate whether NKAP has cellular effects in A549. To do that, the NKAP overexpression vector was generated. As shown in Figure 4.14A, using pcDNA vector carrying MYC-tagged NKAP could efficiently overexpress NKAP in the protein level compared with the control vector in A549. Using BODIPY-C11 with flow cytometry, we found that overexpression of NKAP reduced cellular level of ferroptosis (Figure 4.14C). In addition, we found that overexpression of NKAP increased cell viability compared with the eGFP control vector in A549 cells (Figure 4.14B).

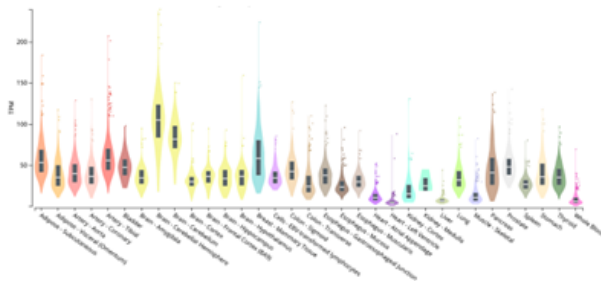
To determine whether downregulation of NKAP that is caused by GRAS1 knockdown is the key factor that contributes to the cellular effects from GRAS1 knockdown, we performed the rescue experiments by NKAP overexpression in GRAS1-lacking A549 cells and see whether it could reverse the effect of ferroptosis and cell viability that was caused by GRAS1 downregulation. Using Annexin V-PI apoptotic assay, we found knockdown of GRAS1 significantly increases the overall apoptotic cell population including the early (Annexin V+/PI-)

and late (Annexin V⁺/PI⁺) apoptotic cells, which led to a big decrease in total survival cell population in the group of GRAS1 knockdown compared with the control (Figure 4.16A). For the rescue experiments, A549 cells were co-transfected with the NKAP overexpression vector as well as the GapmeRs that target GRAS1. As shown in Figure 4.16B, overexpression of NKAP is able to significantly reverse the cell death that was caused by GRAS1 downregulation. Furthermore, we found that overexpression of NKAP strongly reduced the cellular level of DNA damage that was induced by GRAS1 knockdown by measuring γ H2AX protein level (Figure 4.15B). In addition, overexpression of NKAP completely reversed the effect of ferroptosis that was caused by GRAS1 knockdown (Figure 4.14D). These data suggested that NKAP protects A549 cells against cellular effects, including ferroptosis, DNA damage, and cell death, from GRAS1 downregulation.

A



B



C

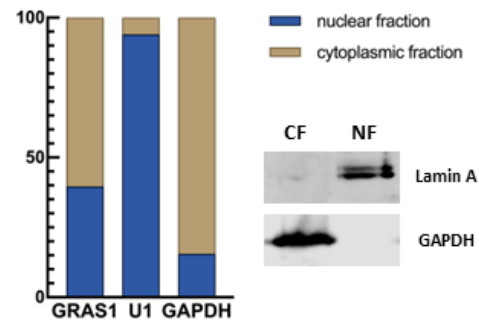


Figure 4.1 GRAS1 is a well-expressed lncRNA located in both nucleus and cytoplasm. (A) Gene track depicting the gene body of GRAS1 with RNA-seq expression region, as well as measured H3K4Me3 and H3K27Ac marks. The data were retrieved from the Encode project and the images were obtained from the UCSC Genome Browser. (B) Expression level of GRAS1 in different tissue. The data were retrieved from the Encode project. (C) RNAs were extracted from fractionated WT A549 followed by analyzed by RT-qPCR. U1 is a nuclear marker and GAPDH is a cytoplasmic marker. Lamin A and GAPDH immunoblots of fractionated cells (cytoplasmic fraction (CF), nuclear fraction (NF)) assessed cell fractionation.

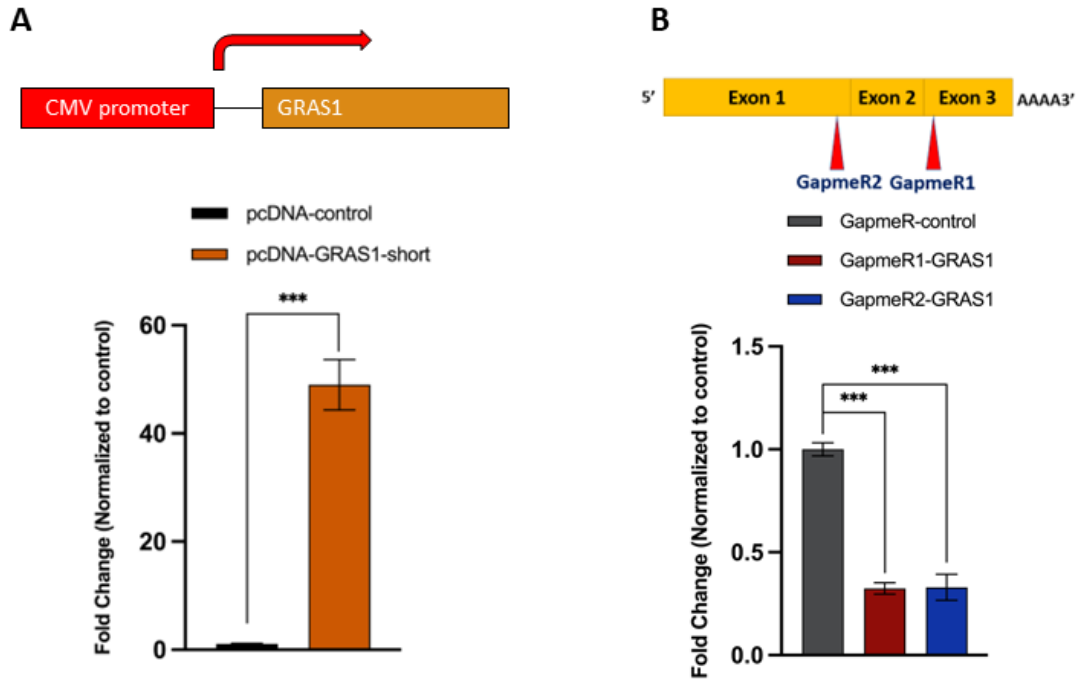


Figure 4.2 GRAS1 can be efficiently overexpressed and downregulated by pcDNA vector and GapmeR respectively. (A) RT-qPCR of GRAS1 overexpression in A549 cells treated with control pcDNA vector compared with pcDNA-GRAS1 vector. Relative fold change calculated based on the geometric mean of GAPDH RNA level. (B) RT-qPCR of GRAS1 knockdown in A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GapmeR2. Relative fold change calculated based on the geometric mean of GAPDH RNA level. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

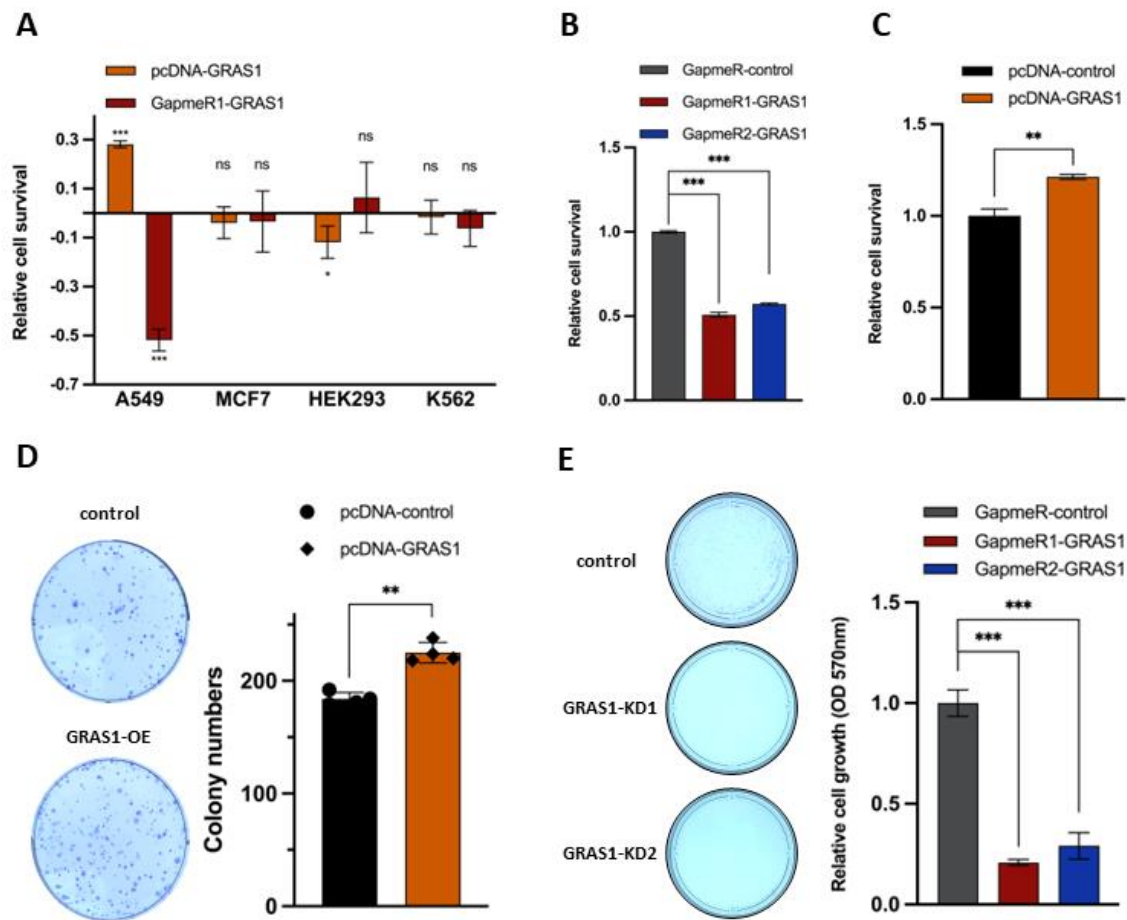


Figure 4.3 GRAS1 is required for lung cancer cell survival. (A) Cell survival was measured by MTT assay on different human cell lines, including A549, MCF7, HEK293, and K562, transfected with control pcDNA vector compared with pcDNA-GRAS1 vector, as well as control GapmeR compared with GRAS1 GapmeR1. Data are presented as fold changes normalized to controls. (B) Cell survival was measured by MTT assay on A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2. (C) Cell survival was measured by MTT assay on A549 transfected with control pcDNA vector compared with pcDNA-GRAS1 vector. (D) Colony formation assay was performed to determine cell viability of A549 cells transfected with control pcDNA vector compared with pcDNA-GRAS1 vector. Quantification of number of formed colonies per biological sample (n=4). (E) Colony formation assay was performed to determine cell viability of A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2. Shown figures are representative images (left) and average in OD values of cell stain in 3 independent experiments (right). Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

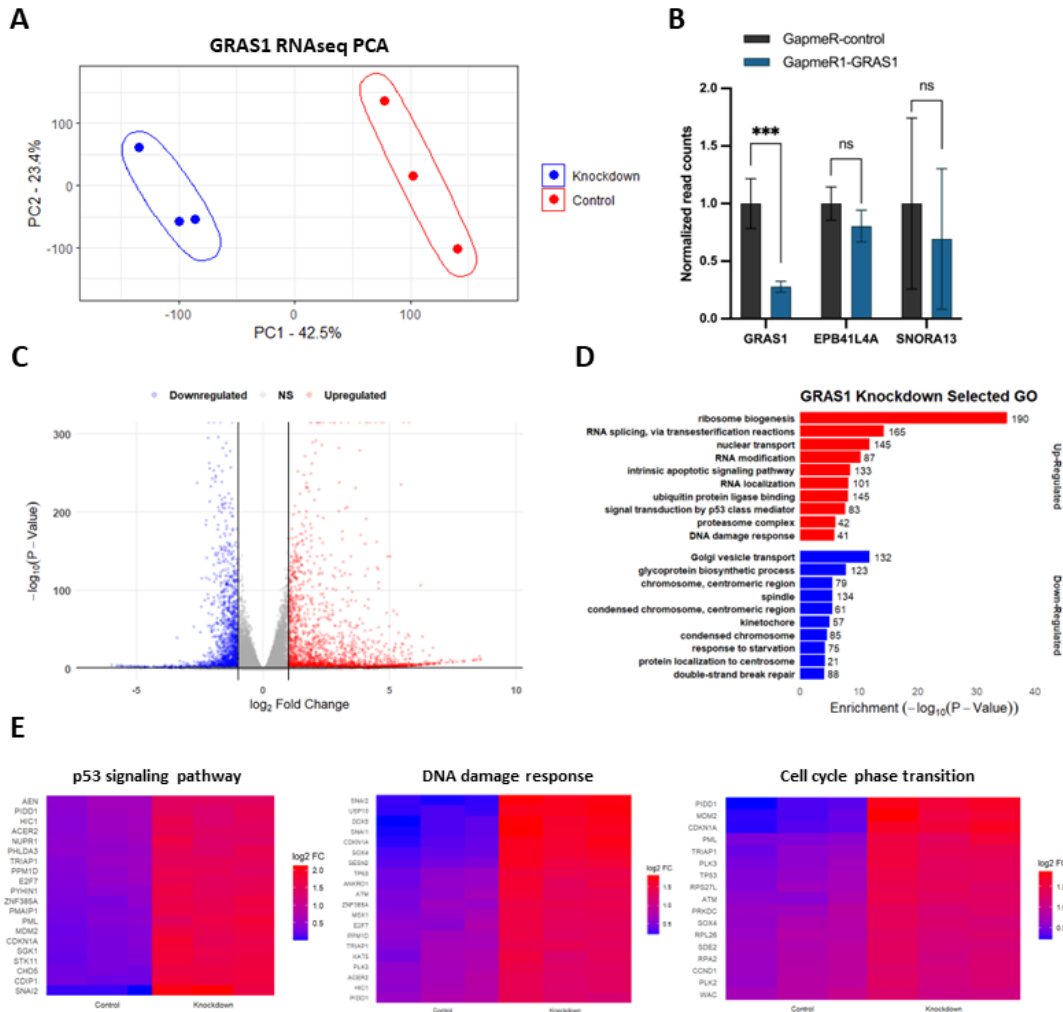


Figure 4.4: GRAS1 affects multiple transcriptomic programs in the lung cancer cell. (A) Principal component analysis (PCA) of the normalized RNAseq data transcripts per million (TPM) of A549 transfected with control GapmeR (red) compared with GRAS1 GapmeR1 (blue). (B) RNA expression of GRAS1, EPB41L4A, and SNORA13 in control and GRAS1 knockdown groups based on read counts in RNA sequencing. (C) Volcano plot representation of differential expression analysis of genes in the control GapmeR versus GRAS1 GapmeR1 comparison for the A549 cells RNA-seq data sets. Red and blue points mark the genes with significantly increased or decreased expression respectively in control GapmeR compared to GRAS1 GapmeR1 samples. The x-axis shows $\log_2$ fold-changes in expression and the y-axis the log odds of a gene being differentially expressed. (D) Gene Ontology enrichment analysis of biological processes for up- and down-regulated genes between control GapmeR and GRAS1 GapmeR1 transfected A549 cells. (E) Heatmap showing the relative gene expression of 20 genes responsible for (A) the p53 signaling pathway (B) the DNA damage response (C) mitotic G1/S transition according to the results of RNA-seq. The horizontal axis represents $\log_2$ fold change. The vertical axis represents the genes. A bright red color indicates stronger up-regulation, and a bright blue color indicates stronger down-regulation in expression. Error bars represent SD. (ns) $P > 0.05$; (***) $P < 0.001$ by two-tailed Student's t-test.

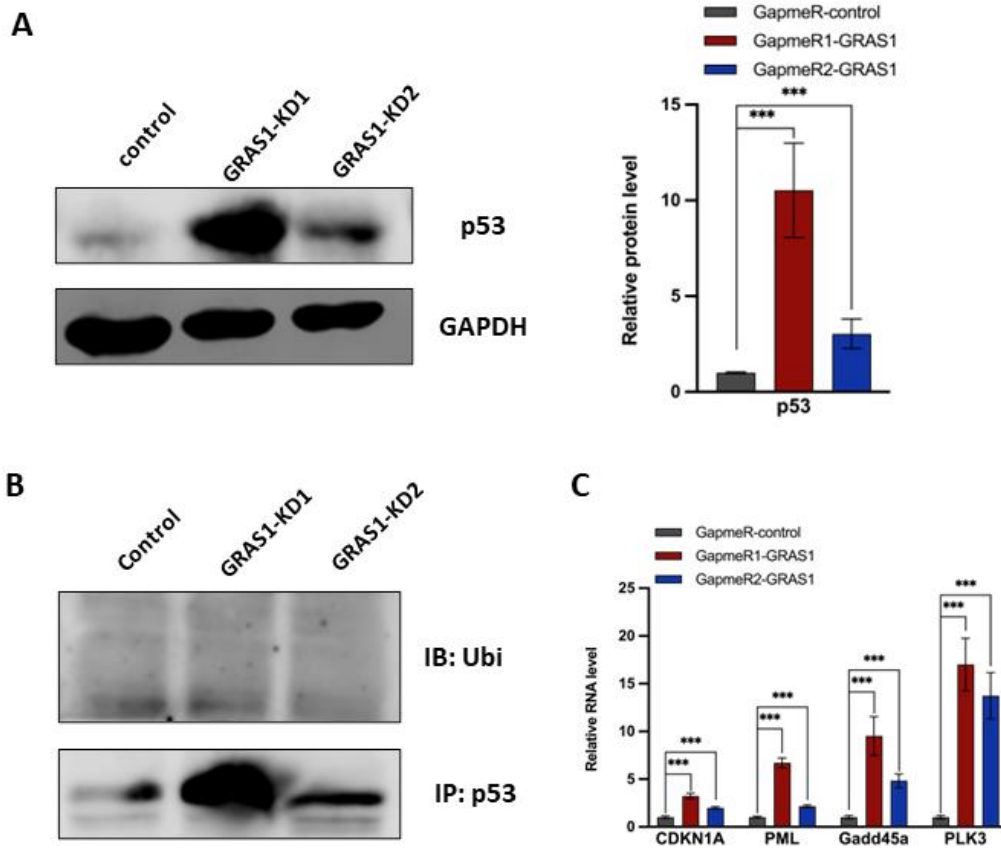


Figure 4.5 Knockdown of GRAS1 activates p53 signaling pathway. (A) Immunoblotting probed for p53 level in A549 comparing control GapmeR with GRAS1 GapmeR1 or GRAS1 GapmeR2 (left). Quantification of p53 expression by normalizing to GAPDH (right). (B) A549 cells were transfected with GapmeR1 or GapmeR2 against GRAS1 along with GapmeR control for 48 h, then p53 was pulled down by anti-p53, and p53 ubiquitination was detected by blotting anti-ubiquitin. (C) mRNA expression of p21, PML, Gadd45a, and PLK3 in A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2 by RT-qPCR. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

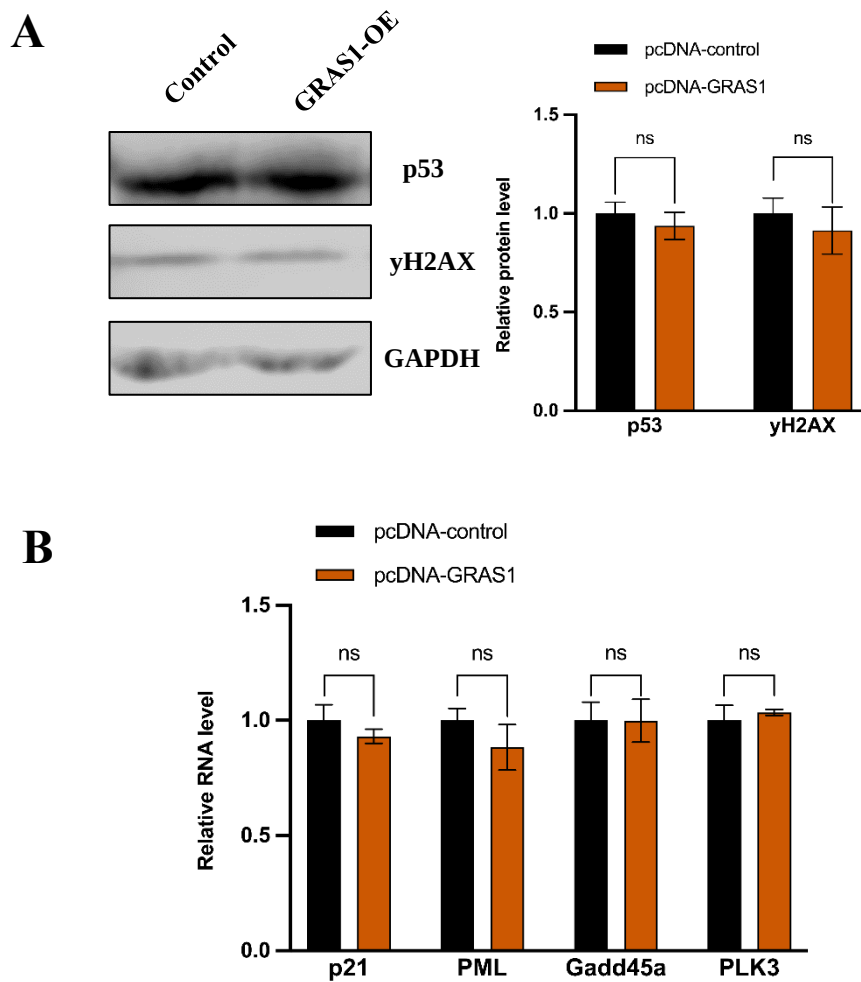
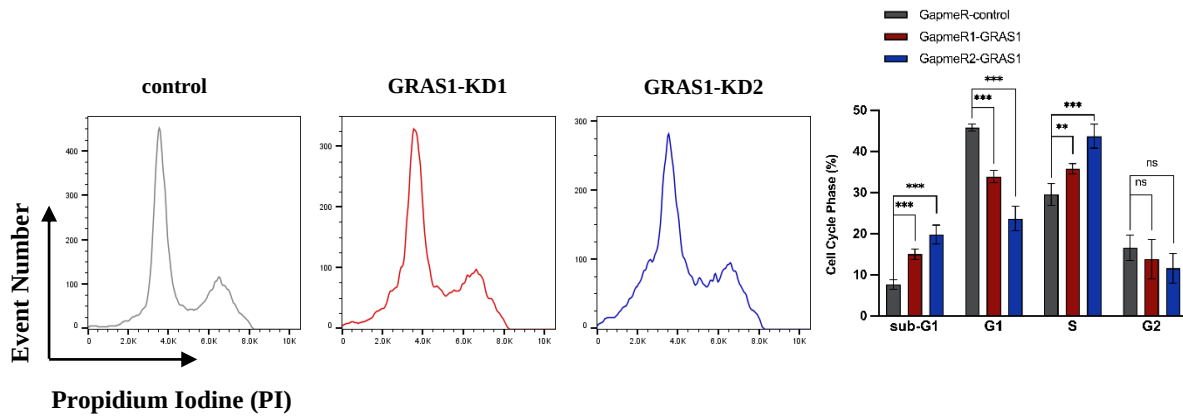


Figure 4.6 Overexpression of GRAS1 induces DNA damage and activates p53 signaling pathway. (A) Immunoblotting probed for p53 level in A549 comparing control vector with GRAS1 overexpression vector (left). Quantification of p53 expression by normalizing to GAPDH (right). (B) mRNA expression of p21, PML, Gadd45a, and PLK3 in A549 cells treated with control vector compared with GRAS1 overexpression vector by qPCR. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

A



B

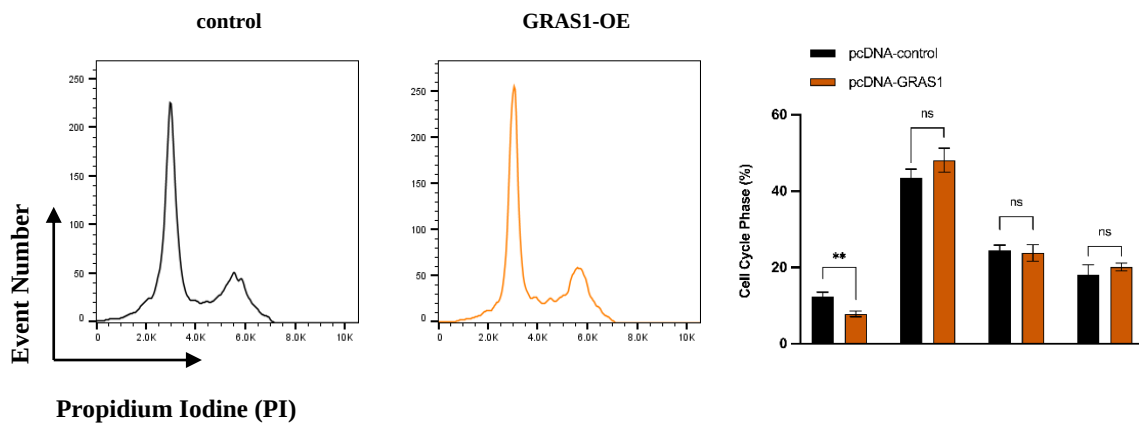


Figure 4.7: GRAS1 affects cell cycle distribution. (A) Knockdown of GRAS1 affects G1/S phase transition from cell cycle analysis. Representative images of cell cycle analysis by PI staining under flow cytometry for A549 transfected with control GapmeR, GRAS1 GapmeR1, and GRAS1 GapmeR2 (left). Quantification of the proportion of cell cycle phase in 3 independent experiments (right). (B) Representative images of cell cycle analysis by PI staining under flow cytometry for A549 transfected with control vector and GRAS1 overexpression vector (left). Quantification of the proportion of cell cycle phase in 3 independent experiments (right). Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

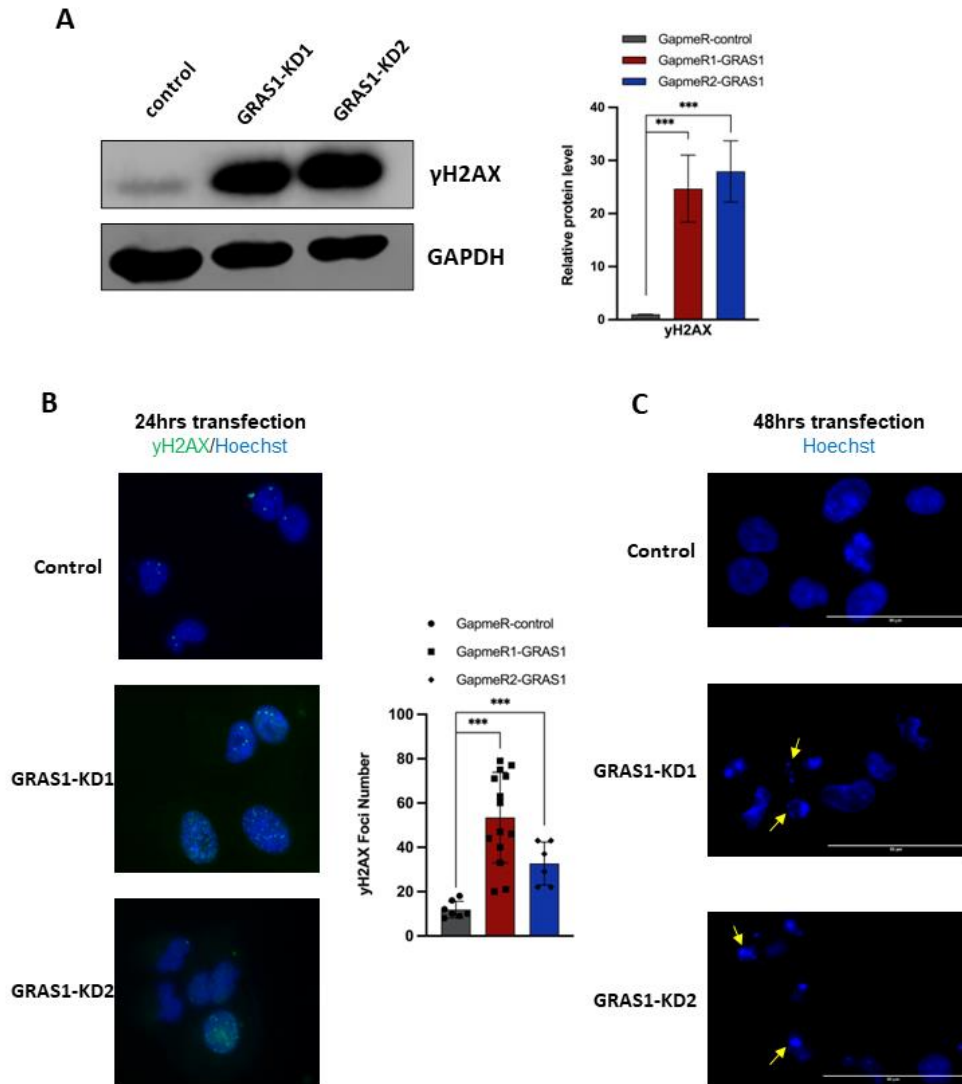


Figure 4.8. Knockdown of GRAS1 induces DNA damage. (A) Immunoblotting probed for γ H2AX level in A549 comparing control GapmeR with GRAS1 GapmeR1 or GRAS1 GapmeR2 (left). Quantification of γ H2AX expression by normalizing to GAPDH (right). (B) γ H2AX foci (green) by immunofluorescence microscopy on A549 after 24hrs transfection with GRAS1 GapmeR1 or GRAS1 GapmeR2 compared to control GapmeR. Hoechst staining in blue. Shown are representative images (left) and the number of foci/cell in 3 independent experiments (right). The scale bar represents 50 μ m. (C) Representative Hoechst DNA staining on A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2. The scale bar represents 50 μ m. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

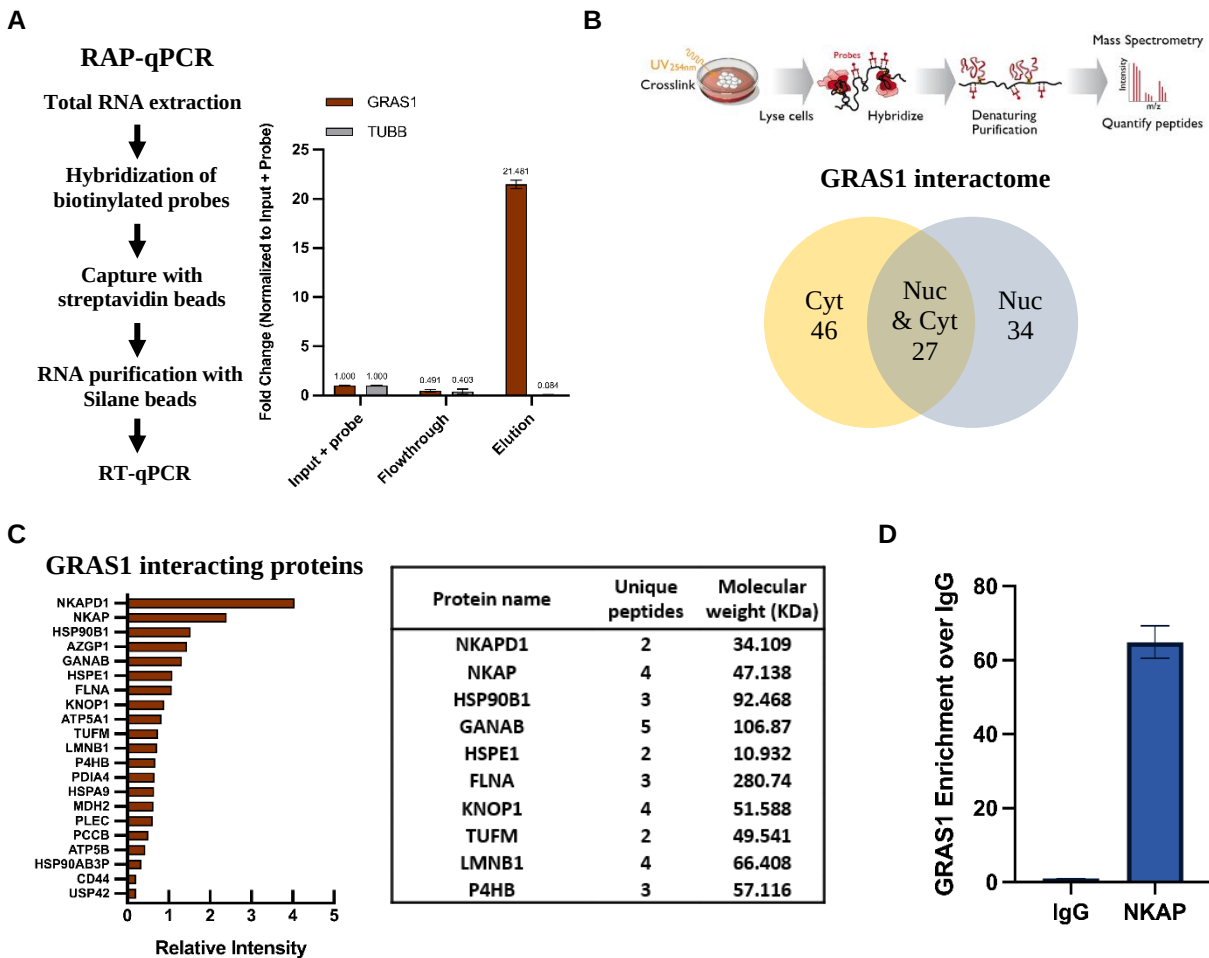


Figure 4.9. GRAS1 directly interacts with NKAP. (A) RNA antisense purification-qPCR experimental scheme to validate biotinylated probes against GRAS1 (left). RT-qPCR on GRAS1 and TUBB in three fractions, including input with probes, flowthrough, and elution. TUBB is a non-specific target. (B) GRAS1 RNA antisense purification-mass spectrometry experimental scheme (top). Cellular components of GRAS1 interactome were shown (bottom). (C) 20 GRAS1-interacting proteins with greatest MS intensity, excluding proteins found in RAP-MS of U1 and Luciferase (left). Table of information on 10 GRAS1-interacting proteins with unique peptides and molecular weight (right). (D) GRAS1 enrichment measured by qPCR after immunoprecipitation on NKAP.

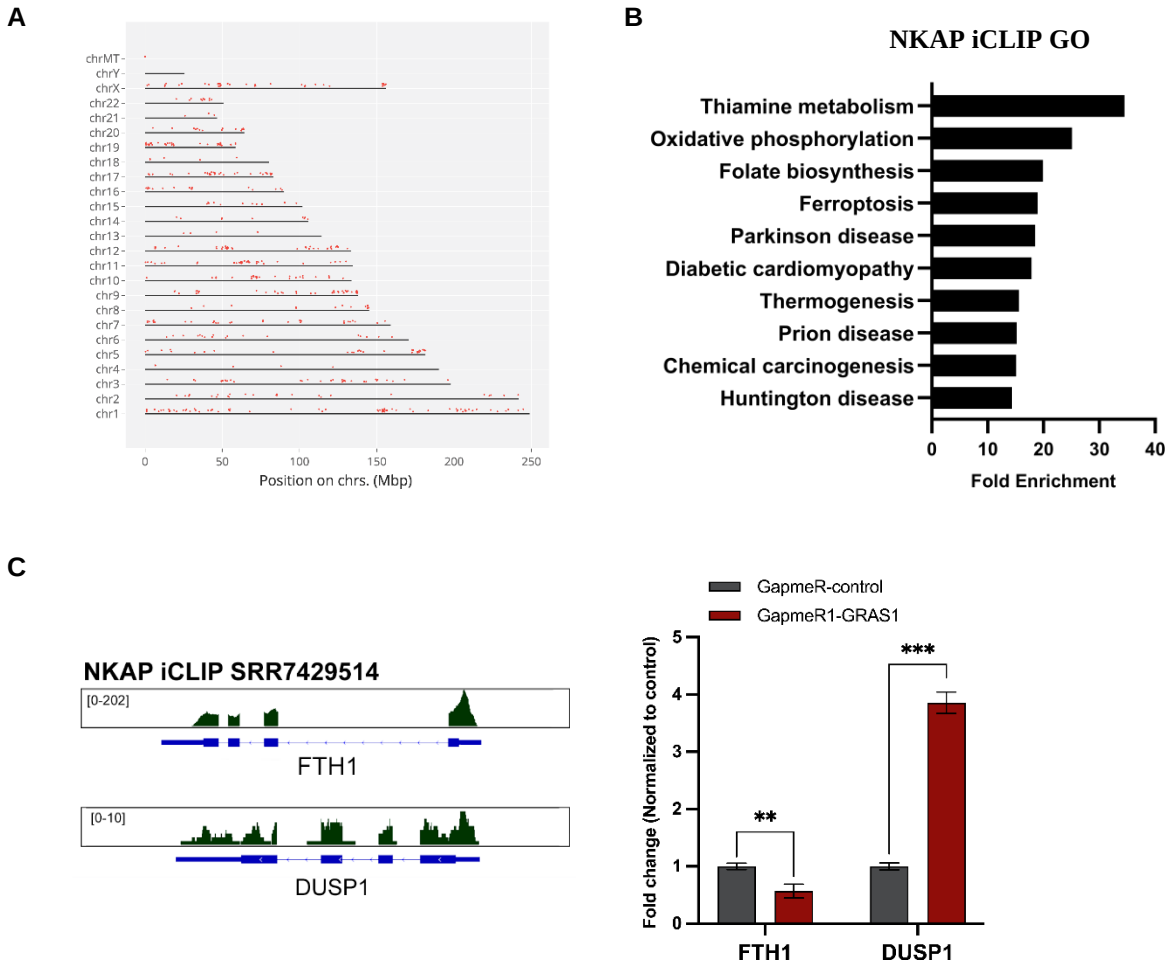


Figure 4.10 NKAP binds with multiple ferroptosis-related RNAs. (A) chromosomal location of the top-1000 NKAP-binding RNAs based on iCLIP signal. (B) Gene Ontology enrichment analysis of biological processes for genes overlapped in NKAP iCLIP and RNA sequencing data of GRAS1 shown in Figure 2. (C) Integrated genome viewer (IGV) browser tracks of the NKAP CLIP peaks spanning the genomic loci of FTH1 and DUSP1 gene (top). RT-qPCR of FTH1 and DUSP1 in A549 treated with GapmeR control or GRAS1 GapmeR1. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

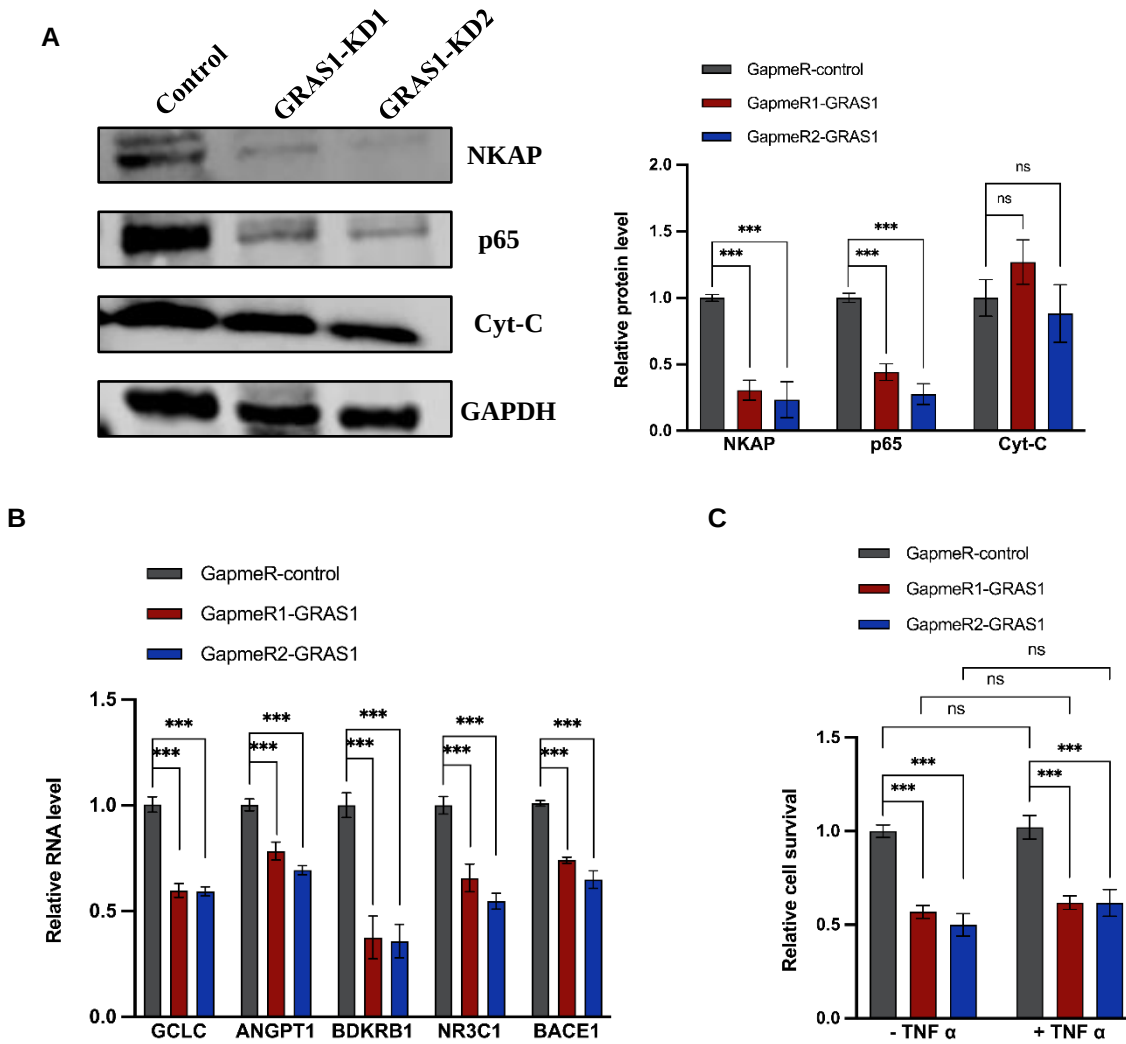


Figure 4.11 Downregulation of GRAS1 decreases NKAP level, which leads to a decrease in NF- κ B signaling pathway. (A) Immunoblotting probed for protein levels of NKAP, p65, and cytochrome-c in A549 comparing control GapmeR with GRAS1 GapmeR1 or GRAS1 GapmeR2 (left). Quantification of expression in NKAP, p65, and cytochrome-c by normalizing to GAPDH (right). (B) mRNA expression of GCLC, ANGPT1, BDKRB1, NR3C1, and BACE1 in A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2 by qPCR. (C) Cell survival was measured by MTT assay on A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2 under condition of with or without 10ng/ml TNF-alpha treatment. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

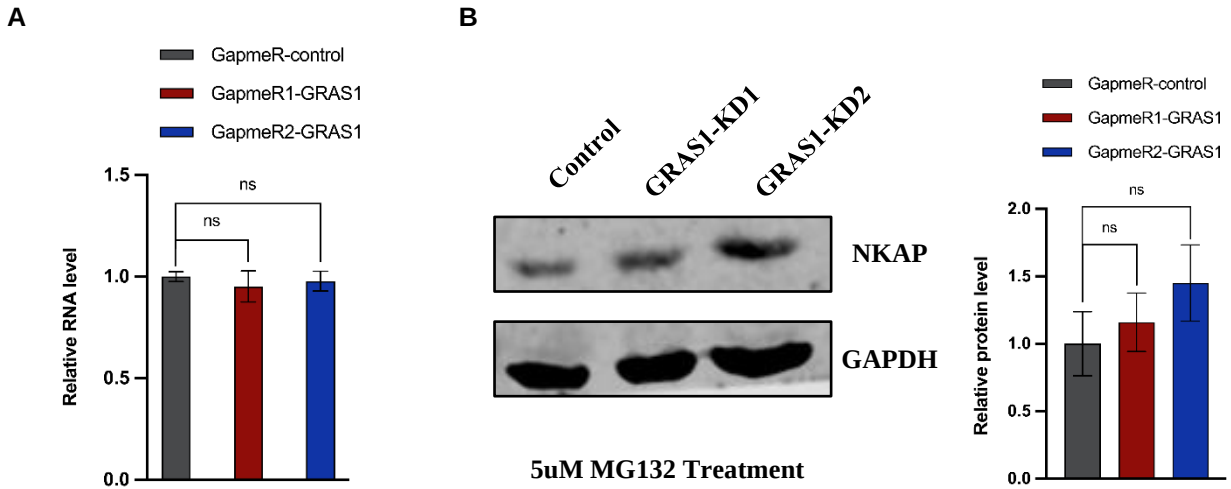


Figure 4.12 GRAS1 post-translationally protects NKAP from ubiquitination. (A) Knockdown of GRAS1 does not affect mRNA level of NKAP. mRNA expression of NKAP in A549 cells treated with control GapmeR compared with GapmeRs that target GRAS1 by qPCR. (B) Immunoblotting probed for protein levels of NKAP in A549 comparing control GapmeR with GRAS1 GapmeR1 or GRAS1 GapmeR2 after 6hrs 5uM MG132 treatment(left). Quantification of expression in NKAP by normalizing to GAPDH (right). Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

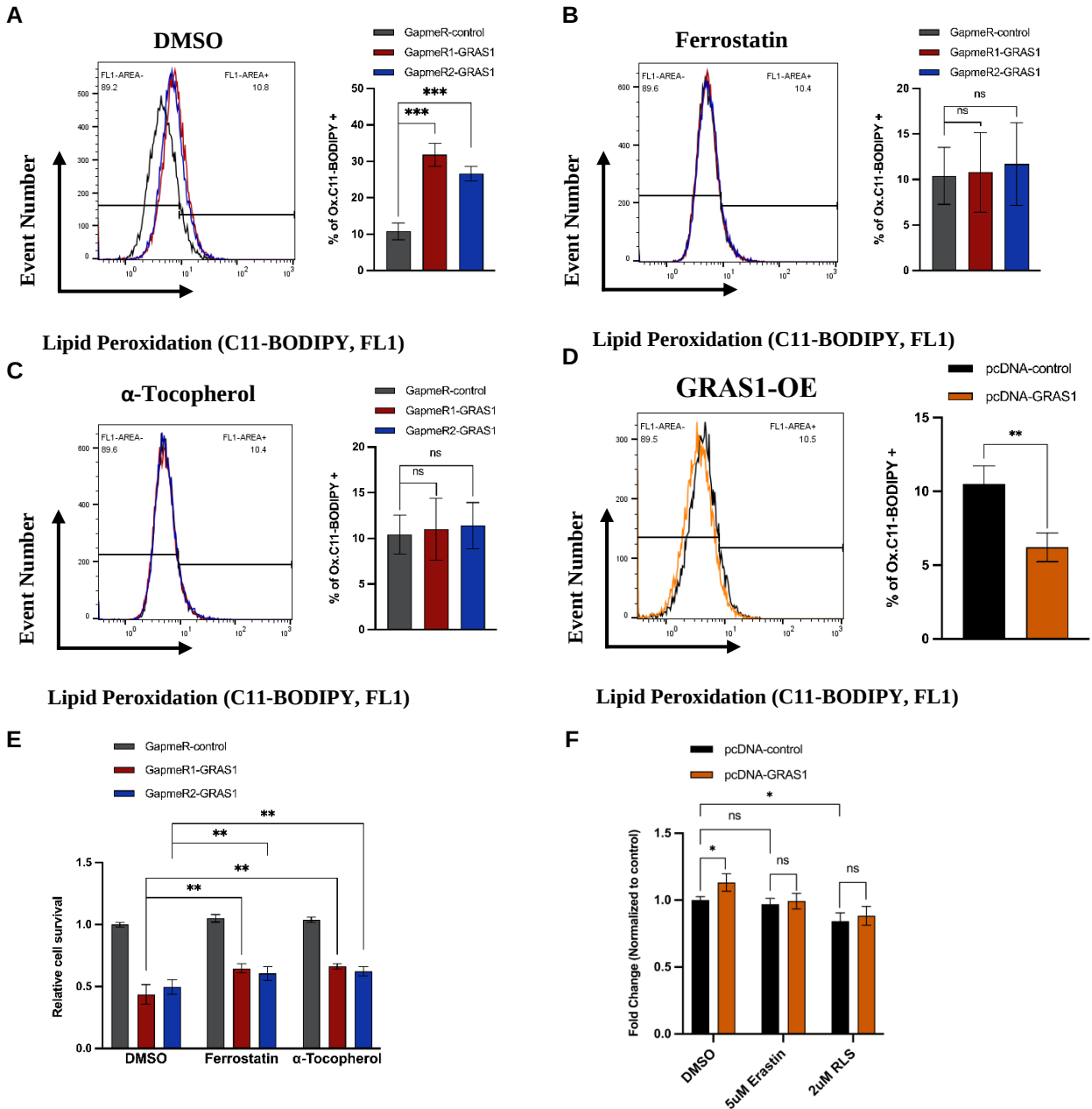


Figure 4.13 GRAS1 affects cellular ferroptosis. Using lipid peroxidation level measured by C11-BODIPY staining under flow cytometry to measure ferroptosis level in A549 cells. (A) Control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2. Representative image (left) and quantification (right) are shown. Lipid peroxidation level influenced by ferroptosis inhibitors, ferrostatin (B) and α -Tocopherol (C). Representative image (left) and quantification (right) are shown. (D) GRAS1 overexpression compared with control vector. (E) Cell survival was measured by MTT assay on A549 cells treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2 under condition of ferrostatin and α -Tocopherol treatment. (F) Cell survival was measured by MTT assay on A549 cells with GRAS1 overexpression compared with the control under condition of Erastin and RLS. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

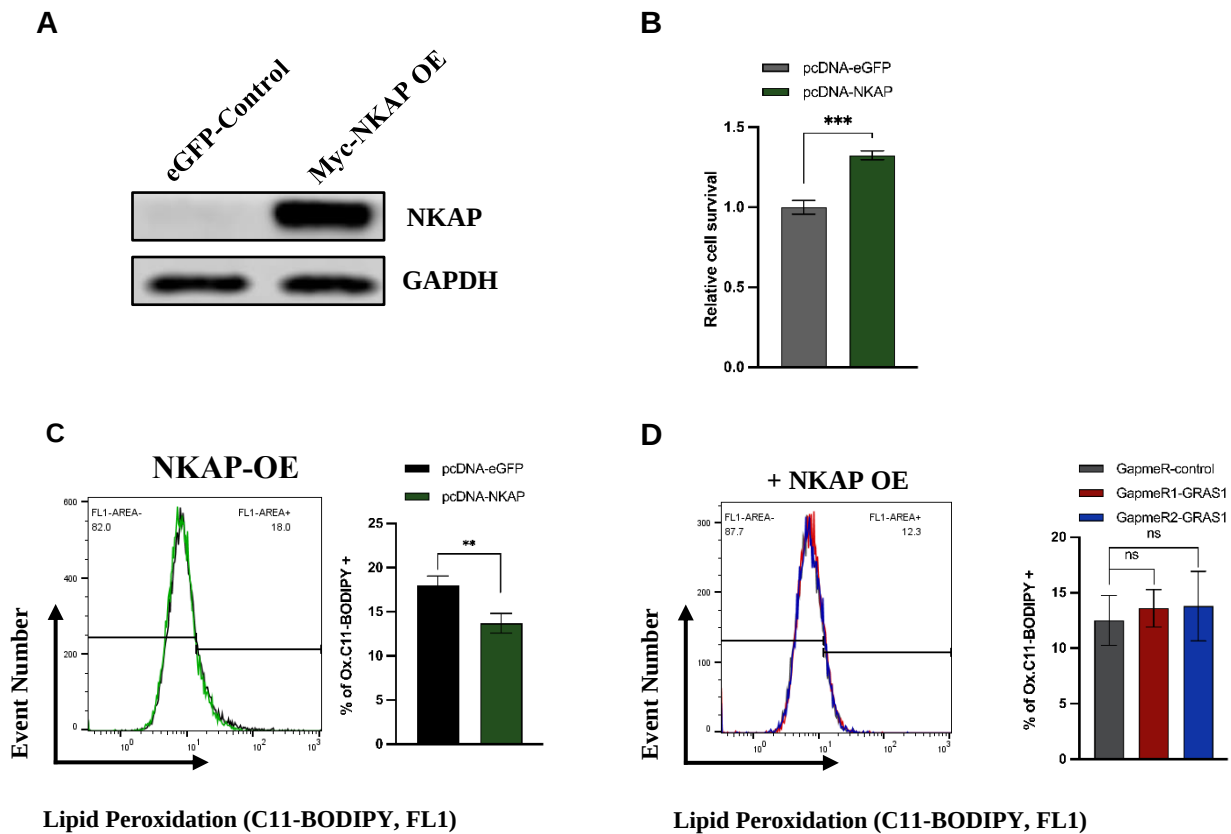
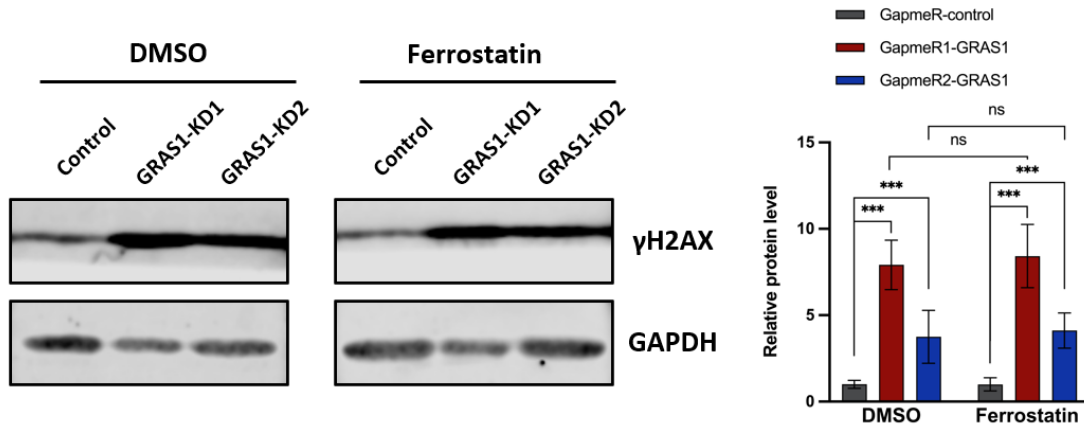


Figure 4.14 Overexpression of NKAP increase cell viability and protects GRAS1-lacking cells from ferroptosis. (A) Immunoblotting probed for protein level of NKAP in A549 transfected with eGFP control vector or myc-tagged NKAP vector. (B) Cell survival was measured by MTT assay on A549 cells transfected with eGFP control vector or myc-tagged NKAP vector. (C) Lipid peroxidation level measured by C11-BODIPY staining under flow cytometry in A549 cells transfected with eGFP control vector or myc-tagged NKAP vector. Representative image (left) and quantification (right) are shown. (D) Lipid peroxidation level influenced by overexpression of NKAP in A549 cell with GRAS1 knockdown, assessed by C11-BODIPY using flow cytometry immunolabelling. Representative image (left) and quantification (right) are shown. Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

A



B

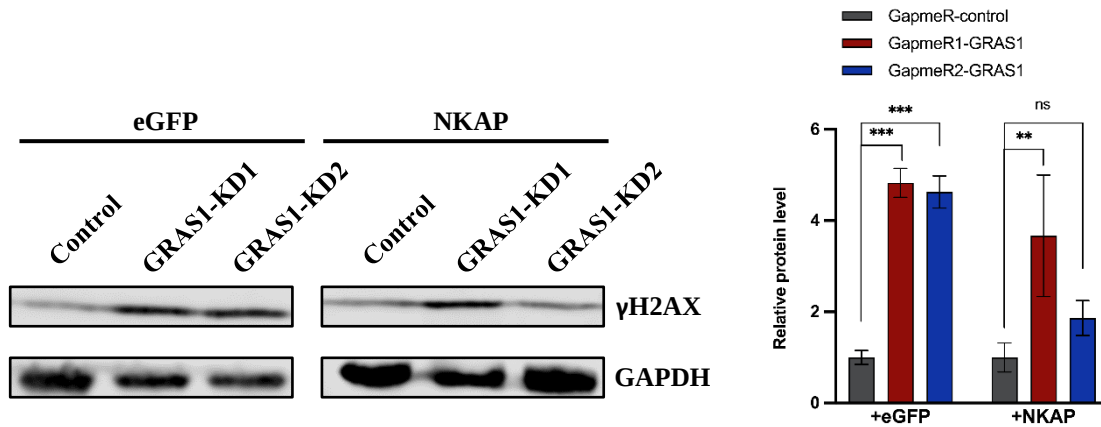


Figure 4.15 Overexpression of NKAP protects GRAS1-lacking A549 from DNA damage.

(A) Immunoblotting probed for protein level of γH2AX in A549 treated with control GapmeR compared with GRAS1 GapmeR1 or GRAS1 GapmeR2 under condition of ferrostatin and α-Tocopherol treatment. (B) Immunoblotting probed for protein level of γH2AX in A549 co-transfected with myc-tagged NKAP vector, comparing control GapmeR with GRAS1 GapmeR1 or GRAS1 GapmeR2 (left). Quantification of expression in γH2AX by normalizing to GAPDH (right). Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

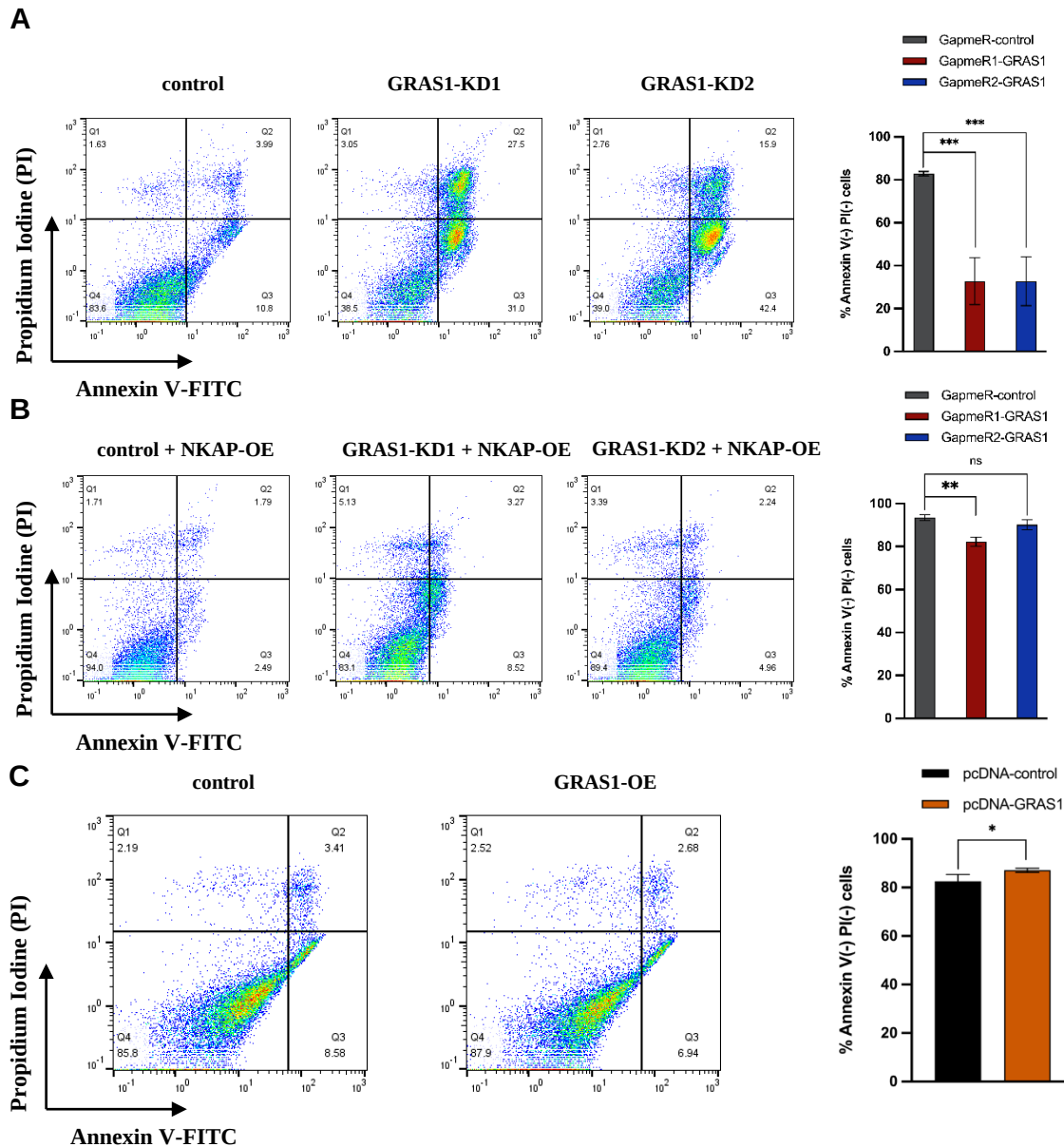


Figure 4.16 Overexpression of NKAP rescues cell death caused by GRAS1 knockdown in A549 cells. Cell survival was assessed by FACS after Annexin V/PI staining. (A) Representative images of A549 with control GapmeR, GRAS1 GapmeR1 or GRAS1 GapmeR2 (left). Quantification of live cells are shown (right). (B) NKAP overexpression increases percentage of live cells from GRAS1 knockdown. Representative images (left), and quantification of live cells are shown (right). (C) GRAS1 protects A549 from cell death. Representative images of A549 with control vector and GRAS1 overexpression vector (left). Quantification of live cells are shown (right). Error bars represent SD. (ns) $P > 0.05$; (*) $P \leq 0.05$; (**) $P \leq 0.01$; (***) $P < 0.001$ by two-tailed Student's t-test.

4.4 Discussion

In this study, we reported that GRAS1, a DNA-damage inducing long non-coding RNA, protects lung cancer cells from ferroptosis. Ectopic expression of GRAS1 by pcDNA overexpression vector increases cell survival in the lung cancer cell line A549. On the contrary, downregulation of GRAS1 by GapmeR knockdown dramatically decreases cell viability in A549. Consistent with the effects of cell survival we observed from MTT assay, overexpression of GRAS1 reduces the level of cell death, while knockdown of the lncRNA leads to a significant increase in apoptotic level in A549 using the Annexin V apoptosis assay. In addition, GRAS1 knockdown had a highly increasing cellular level of chromosome condensation and apoptotic bodies in A549, showing lack of GRAS1 by GapmeR knockdown induces cell death in A549. It is surprising that the effects of GRAS1 on cell viability was only observed in the lung cancer cell line A549 based on our MTT assay cell survival screening using a total 5 different cell lines, suggesting GRAS1 may function specifically in lung cancer cells although more lung cancer cell lines needed to be examined. In addition, in order to address the fact that GRAS1 is critical in the development of lung cancer, effects of GRAS1 on primary lung fibroblast still need to be tested.

Knockdown of GRAS1 dramatically enhances the protein level of the p53, as well as the mRNA level of the p53 downstream target genes in A549. Surprisingly, no significant changes were observed on both p53 level and expression of the p53 downstream target genes when GRAS1 was overexpressed in A549. Mammalian cells usually maintain a low level of p53 under unstressed condition, because p53 has a short half-life as well as it continuously gets ubiquitinated and degraded in cells (Lakin and Jackson 1999). Interestingly, we found downregulation of GRAS1 does not increase the ubiquitination of p53, suggesting that the increased level of p53 may result from a different pathway. In addition, we found GRAS1 affects

cell cycle distribution in A549. Here, we reported that overexpression of GRAS1 decreases while knockdown of GRAS1 increases the sub-G1 phase in the cell cycle. The sub-G1 phase represents the estimation of fragmented DNA content, which is one of the characteristics of cell death (Shu et al. 1997), so the data of cell cycle is consistent with the cell survival assay on GRAS1.

Cellular p53 accumulation can be caused by multiple stressing factors, including DNA damage, metabolic alterations, hypoxia, abnormal oncogenic expression, and cytokine activations (Lakin and Jackson 1999). Here, we reported that lack of GRAS1 by GapmeR knockdown significantly enhances cellular level of DNA damage through quantitatively measuring the DNA damage marker γ H2AX. Hypoxia induces activation of p53 signaling pathway as well as being the most common microenvironmental factor for lung cancer progression (Swinson and O'Byrne 2006). Hypoxia-inducible factor 1- α (HIF1 α) is considered as the master protein in response to hypoxia, and it gets strongly upregulated under the hypoxia condition (Wang et al. 1995, Minet et al. 1999). We found that either GRAS1 overexpression or GRAS1 knockdown does not affect the protein level of HIF1 α in A549. Although we did not examine every stressing factor that is able to contribute to p53 activation, DNA damage is certainly one of the reasons that causes an increase level in p53 when GRAS1 is downregulated in A549.

GRAS1 has been shown to be dysregulated in several types of cancer (Delshad et al. 2019, Wang et al. 2020, Bin et al. 2021). A few functional studies have demonstrated that GRAS1 affects multiple biological processes, including cellular metabolism, cell cycle (Liao et al. 2019, Cui et al. 2020), suggesting GRAS1 may affect multiple transcriptomic programs. It is surprising that downregulation of GRAS1 leads to a large number of differentially expressed genes in A549 based on RNA sequencing. Furthermore, GO analysis included multiple pathways

related to cell survival, including p53 activation, DNA damage response, and cell cycle distribution as confirmed by biochemical assays.

Several studies have shown lncRNAs can act as scaffolds to recruit functional proteins and subsequently affect biological processes (Roberts et al. 2014). Unlike regular RNA pulldown assay, RAP-MS is highly selective tool for identifying specific and direct protein interacting partners with RNA molecules, which was previously used successfully to identify the XIST non-coding RNA interaction with SPEN/SHARP in initiation of X-chromosome silencing (McHugh et al. 2015). We identified 18 proteins that specifically interact with GRAS1 using the RAP-MS, including multiple mitochondrial proteins involved in the cellular metabolism (GANAB, TUFM, P4HB, PDIA4, MDH2, PCCB, ATP5B), a few heat shock proteins (HSPE1, HSPA4, HSP90) and one protein involved in cell growth signaling pathway (NKAP). NKAP is one of the proteins with highest MS intensity from the RAP-MS on GRAS1, and their interaction was validated by RNA immunoprecipitation. NKAP is a multifunctional protein and is essential for growth and development processes in multiple highly proliferative cell types including spermatozoa, T cells, and hematopoietic stem cells. (Pajerowski et al. 2009, Shapiro et al. 2019, Okuda et al. 2015). Surprisingly, we found that the knockdown of GRAS1 reduces the protein level of NKAP, while the mRNA level of NKAP does not change. This data indicates that GRAS1 affects NKAP post-translationally. Our data gives an insight that GRAS1 may protect NKAP from ubiquitination, but further studies are required to investigate the detailed mechanism on how the interaction between GRAS1 and NKAP regulates NKAP stability.

NKAP is involved in multiple biological processes, including NF- κ B pathway (Liu et al, 2021) and ferroptosis (Sun et al, 2022). Our data show downregulation of NKAP reduces the protein level of p65 as well as the activity of NF- κ B pathway. However, activating the NF- κ B

pathway by TNF- α does not rescue the cell death caused by GRAS1 knockdown, suggesting that NF-kB pathway does not contribute to the effect of GRAS1 on cell viability in A549.

Furthermore, we found knockdown of GRAS1 induces cellular ferroptosis. Ferroptosis is a type of cell death caused by excessive cellular iron that is distinct from apoptosis, necrosis, and autophagy, and it can be regulated by p53 (Dixon and Stockwell. 2019, Xu et al. 2019). Using ferroptosis inhibitors, including ferrostatin and α -tocopherol, can reverse the increased level of ferroptosis as well as partially rescue cell death that is caused by GRAS1 downregulation. Here, we found that overexpression of NKAP significantly reduces cellular level of ferroptosis and increases cell viability in A549. In addition, A549 cells that were lacking GRAS1 can significantly get rescued in cell viability by ectopic expression of NKAP, suggesting NKAP is the key factor that links to the decrease in cell survival caused by GRAS1 downregulation in A549. Ferroptosis leads to accumulation of cellular reactive oxygen species (Xu et al. 2019), however, the connections between ferroptosis and DNA damage is still poorly understood. Interestingly, we found overexpression of NKAP reverses the increased level of γ H2AX caused by GRAS1 knockdown. However, no significant difference was observed in the protein level of γ H2AX in A549 with GRAS1 knockdown under the treatment of ferroptosis inhibitors, suggesting the increased level of γ H2AX by GRAS1 knockdown is unrelated to ferroptosis. Further studies are needed to investigate whether NKAP regulates different pathways that involve cellular DNA damage and ferroptosis.

Reference:

1. Veena, V. S., George, P. S., Jayasree, K., & Sujathan, K. (2013). Lung cancer is one of the leading cause of cancer related deaths for people of both sexes worldwide and sputum

cytology has been accepted as one of the best non-invasive method for lung cancer diagnosis but this method is not accepted as a screening techn. *Proc. Int. J. Sci. Res. Publ.*, 2(12), 1-6.

2. de Groot, P. M., Wu, C. C., Carter, B. W., & Munden, R. F. (2018). The epidemiology of lung cancer. *Translational lung cancer research*, 7(3), 220.
3. World Health Organization. (2020). Cancer. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/cancer>
4. National Cancer Institute. (2021). Non-Small Cell Lung Cancer Treatment (PDQ®)–Patient Version. Retrieved from <https://www.cancer.gov/types/lung/patient/non-small-cell-lung-treatment-pdq>
5. Mayo Clinic. (2021). Lung cancer. Retrieved from <https://www.mayoclinic.org/diseases-conditions/lung-cancer/diagnosis-treatment/drc-20374629>
6. Kazemzadeh, M., Safaralizadeh, R., & Orang, A. V. (2015). LncRNAs: emerging players in gene regulation and disease pathogenesis. *Journal of genetics*, 94, 771-784.
7. Wang, K.C., & Chang, H.Y. (2011) "Molecular mechanisms of long noncoding RNAs." *Molecular Cell* 43(6), 904-914.
8. Wang, M., Zheng, S., Li, X., Ding, Y., Zhang, M., Lin, L., Xu, H., Cheng, Y., Zhang, X., Xu, H. and Li, S., 2020. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network identified lncRNA EPB41L4A-AS1 as a potential biomarker in non-small cell lung cancer. *Frontiers in Genetics*, 11, p.511676.
9. Delshad, E., Shamsabadi, F. T., Bahramian, S., Mehravar, F., Maghsoudi, H., & Shafiee, M. (2019). In silico identification of novel lncRNAs with a potential role in diagnosis of gastric cancer. *Journal of Biomolecular Structure and Dynamics*.

10. Bin, J., Nie, S., Tang, Z., Kang, A., Fu, Z., Hu, Y., Liao, Q., Xiong, W., Zhou, Y., Tang, Y. and Jiang, J., 2021. Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating the Rho/ROCK pathway in colorectal cancer. *Journal of cellular physiology*, 236(1), pp.523-535.
11. Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., Wang, Y., Wang, S., Zhu, Y., Chen, D. and Xie, W., 2019. LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. *EBioMedicine*, 41, pp.200-213.
12. Zhu, Y., Liu, Q., Liao, M., Diao, L., Wu, T., Liao, W., Wang, Z., Li, B., Zhang, S., Wang, S. and Xie, W., 2019. Overexpression of lncRNA EPB41L4A-AS1 induces metabolic reprogramming in trophoblast cells and placenta tissue of miscarriage. *Molecular Therapy-Nucleic Acids*, 18, pp.518-532.
13. Cui, P., Zhao, X., Liu, J., Chen, X., Gao, Y., Tao, K., Wang, C. and Zhang, X., 2020. miR-146a interacting with lncRNA EPB41L4A-AS1 and lncRNA SNHG7 inhibits proliferation of bone marrow-derived mesenchymal stem cells. *Journal of Cellular Physiology*, 235(4), pp.3292-3308.
14. Chen D, Li Z, Yang Q, Zhang J, Zhai Z, Shu HB. Identification of a nuclear protein that promotes NF-kappaB activation. *Biochem Biophys Res Commun*. 2003 Oct 24;310(3):720-4. doi: 10.1016/j.bbrc.2003.09.074. PMID: 14550261.
15. Burgute BD, Peche VS, Steckelberg AL, Glöckner G, Gaßen B, Gehring NH, Noegel AA. NKAP is a novel RS-related protein that interacts with RNA and RNA binding proteins. *Nucleic Acids Res*. 2014 Mar;42(5):3177-93. doi: 10.1093/nar/gkt1311. Epub 2013 Dec 17. PMID: 24353314; PMCID: PMC3950704.

16. Wei Y, Wang Y, Zang A, Wang Z, Fang G, Hong D. MiR-4766-5p Inhibits The Development And Progression Of Gastric Cancer By Targeting NKAP. *Onco Targets Ther.* 2019 Oct 16;12:8525-8536. doi: 10.2147/OTT.S220234. PMID: 31802890; PMCID: PMC6801498.
17. Shu W, Liu G, Dai Y, Feng A, Chen Z, Han J, Li X. The oncogenic role of NKAP in the growth and invasion of colon cancer cells. *Oncol Rep.* 2019 Nov;42(5):2130-2138. doi: 10.3892/or.2019.7322. Epub 2019 Sep 18. PMID: 31545474.
18. Ma Q, Hou L, Gao X, Yan K. NKAP promotes renal cell carcinoma growth via AKT/mTOR signalling pathway. *Cell Biochem Funct.* 2020 Jul;38(5):574-581. doi: 10.1002/cbf.3508. Epub 2020 Feb 7. PMID: 32032976.
19. Sun S, Gao T, Pang B, Su X, Guo C, Zhang R, Pang Q. RNA binding protein NKAP protects glioblastoma cells from ferroptosis by promoting SLC7A11 mRNA splicing in an m6A-dependent manner. *Cell Death Dis.* 2022 Jan 21;13(1):73. doi: 10.1038/s41419-022-04524-2. PMID: 35064112; PMCID: PMC8783023.
20. Li, T., Chen, L., Cheng, J. et al. SUMOylated NKAP is essential for chromosome alignment by anchoring CENP-E to kinetochores. *Nat Commun* 7, 12969 (2016).
21. Dixon, S. J., & Stockwell, B. R. (2019). The hallmarks of ferroptosis. *Annual Review of Cancer Biology*, 3(1), 35–54.
22. Xu, T., Ding, W., Ji, X., Ao, X., Liu, Y., Yu, W. and Wang, J., 2019. Molecular mechanisms of ferroptosis and its role in cancer therapy. *Journal of cellular and molecular medicine*, 23(8), pp.4900-4912.
23. Li, J., Cao, F., Yin, H.L., Huang, Z.J., Lin, Z.T., Mao, N., Sun, B. and Wang, G., 2020. Ferroptosis: past, present and future. *Cell death & disease*, 11(2), p.88.

24. Ursini, F., & Maiorino, M. (2020). Lipid peroxidation and ferroptosis: The role of GSH and GPX4. *Free Radical Biology and Medicine*, 152, 175–185.
25. Chen, X., Kang, R., Kroemer, G. and Tang, D., 2021. Broadening horizons: the role of ferroptosis in cancer. *Nature reviews Clinical oncology*, 18(5), pp.280-296.
26. Toné, S., Sugimoto, K., Tanda, K., Suda, T., Uehira, K., Kanouchi, H., Samejima, K., Minatogawa, Y., & Earnshaw, W. C. (2007). Three distinct stages of apoptotic nuclear condensation revealed by time-lapse imaging, biochemical and electron microscopy analysis of cell-free apoptosis. *Experimental cell research*, 313(16), 3635–3644.
27. Kastan, M. B., Canman, C. E., & Leonard, C. J. (1995). P53, cell cycle control and apoptosis: implications for cancer. *Cancer and Metastasis reviews*, 14, 3-15.
28. Lakin, N., Jackson, S. Regulation of p53 in response to DNA damage. *Oncogene* 18, 7644–7655 (1999).
29. Mah, L. J., El-Osta, A., & Karagiannis, T. C. (2010). γ H2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia*, 24(4), 679-686.
30. Liu, J., Chen, S. J., Hsu, S. W., Zhang, J., Li, J. M., Yang, D. C., Gu, S., Pinkerton, K. E., & Chen, C. H. (2021). MARCKS cooperates with NKAP to activate NF- κ B signaling in smoke-related lung cancer. *Theranostics*, 11(9), 4122–4136.
31. Tang, D., Tao, D., Fang, Y., Deng, C., Xu, Q., & Zhou, J. (2017). TNF-Alpha Promotes Invasion and Metastasis via NF-Kappa B Pathway in Oral Squamous Cell Carcinoma. *Medical science monitor basic research*, 23, 141–149
32. Martinez, A. M., Kim, A., & Yang, W. S. (2020). Detection of Ferroptosis by BODIPY™ 581/591 C11. *Methods in molecular biology* (Clifton, N.J.), 2108, 125–130.

33. Lakin, N. D., & Jackson, S. P. (1999). Regulation of p53 in response to DNA damage. *Oncogene*, 18(53), 7644-7655.
34. Swinson, D. E. B., & O'Byrne, K. J. (2006). Interactions between hypoxia and epidermal growth factor receptor in non-small-cell lung cancer. *Clinical lung cancer*, 7(4), 250-256.
35. Wang, G. L., Jiang, B. H., Rue, E. A., & Semenza, G. L. (1995). Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension. *Proceedings of the national academy of sciences*, 92(12), 5510-5514.
36. Minet, E., Ernest, I., Michel, G., Roland, I., Remacle, J., Raes, M., & Michiels, C. (1999). HIF1A gene transcription is dependent on a core promoter sequence encompassing activating and inhibiting sequences located upstream from the transcription initiation site and cis elements located within the 5' UTR. *Biochemical and biophysical research communications*, 261(2), 534-540.
37. Roberts, T. C., Morris, K. V., & Weinberg, M. S. (2014). Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*, 9(1), 13-20.
38. Pajerowski, A. G., Nguyen, C., Aghajanian, H., Shapiro, M. J., & Shapiro, V. S. (2009). NKAP is a transcriptional repressor of notch signaling and is required for T cell development. *Immunity*, 30(5), 696-707.
39. Shapiro, M. J., Anderson, J., Lehrke, M. J., Chen, M., Nelson Holte, M., & Shapiro, V. S. (2019). NKAP regulates senescence and cell death pathways in hematopoietic progenitors. *Frontiers in Cell and Developmental Biology*, 7, 214.
40. Okuda, H., Kiuchi, H., Takao, T., Miyagawa, Y., Tsujimura, A., Nonomura, N., Miyata, H., Okabe, M., Ikawa, M., Kawakami, Y. and Goshima, N., 2015. A novel transcriptional factor

Nkapl is a germ cell-specific suppressor of Notch signaling and is indispensable for spermatogenesis. *PLoS One*, 10(4), p.e0124293.

4.6 Acknowledgement

Chapter 4, in full, is currently being submitted for publication. The authors in the publication: Tong Su, Lingbo Kong, Darin Cheung, Jonathan Zhu, Nichelle Camden, Wei-Han Chou, Calvin Huang, and Colleen McHugh. The dissertation author is the primary author of this material. We thank Brian Zid (University of California San Diego) for assistance with confocal microscopy and flow cytometry. This work was supported by Startup Funding from university of California San Diego, NCI Cancer Training grant T32 CA009523 (Tong Su), University of California Cancer Research Coordinating Committee grant support C21CR2119, Undergraduate Summer Research Award (Lingbo Kong, Calvin Huang).

Chapter 5: Material and Methods

5.1 Cell lines and culture conditions

Human lung adenocarcinoma A549 cell line (ATCC No. CCL-185), human colorectal carcinoma HCT116 cell line (ATCC No. CCL-247), human breast adenocarcinoma MCF7 cell line (ATCC No. HTB-22), human kidney embryonic HEK293 cell line (ATCC No. CRL-1573), and human chronic myeloid leukemia K562 cell line (ATCC, No. CRL-3344) were purchased from ATCC. A549, MCF7, and HEK293 cells were cultured in Eagle's Minimal Essential Medium (EMEM) (Quality Biological, USA); K562 cell was cultured in Iscove's Modified Dulbecco's Medium (IMDM) (Gibco, Life Technologies, USA); HCT116 cell was cultured in McCoy's 5A Modified Medium (Gibco, Life Technologies, USA). All mediums were supplemented with 10% fetal bovine serum (Gibco, Life Technologies, USA), 1% additional L-Glutamine (Corning, USA). All cells were grown in a cell culture incubator at 37 C and 5% CO₂. All cells were tested negative for Mycoplasma.

5.2 Plasmid generation and transfection

Gene inserts were generated by K562 cDNA via PCR. Entry vectors of the Gateway cloning system were made using pENTR/D-TOPO Cloning Kit (2402649, Invitrogen, CA, USA). Once the entry vectors were generated, the gene inserts were swapped into the pcDNA3.2 destination vectors by Gateway LR Clonase Enzyme Mix (11791011, ThermoFisher Scientific, USA). For the plasmid transfection, cells were seeded with 0.2M cells/well in 12-well plates overnight before each treatment. Then the cells in each well were transfected with 1ug of plasmid using

3.5ul of Lipofectamine 2000 (ThermoFisher Scientific, USA) in Opti-MEM I reduced serum medium (31985062, Life Technologies, USA) following the manufacturer's suggested protocol.

5.3 Gene knockdown

For the siRNA knockdown, cells were seeded with 0.2M cells/well in 12-well plates overnight before each treatment. Then cells in each well were transfected with 50 uM of GRAS1-siRNA (Dharmacon, USA) with the GAS5-siRNA (D-001310-01-05, Dharmacon, USA) as a positive control and Non-targeting siRNA (D-001320-01-05, Dharmacon, USA) using 3.5ul of Lipofectamine 2000 in Opti-MEM I reduced serum medium following the manufacturer's suggested protocol. siRNA target sequence is shown in Table 1. For the GapmeR knockdown, cells in each well were transfected with 50 uM of each lncRNA-target GapmeR (Qiagen) with Negative Control GapmeR A (30300019-2, Qiagen) using 3.5ul of Lipofectamine 2000 in Opti-MEM I reduced serum medium following the manufacturer's suggested protocol. GapmeR target sequence is shown in Table 2.

5.4 RNA, DNA Isolation and qRT- PCR

Total RNA from all cell treatments was isolated using QIAshredder and RNeasy Mini kit (Qiagen) then cleaned up with RNA clean & concentrator kit (Zymo Research) for a purer product. The total RNA of each cell fraction from the fractionation experiments were isolated using the RNA cleanup kit. DNA isolation was performed using plasmid miniprep kit (Qiagen). All isolations were performed according to the manufacturer's protocol. 1.5 ug of purified RNA was used for each reverse transcription reaction with 1ul of Superscript III Reverse Transcriptase (18080085, Invitrogen, USA). The qPCR was done with Taq polymerase (New England Biolabs), ROX Reference Dye (Invitrogen, USA), SYBR Green Dye (Life Technologies, USA).

Samples were repeated in triplicates, data were normalized to IgG for immunoprecipitation experiments, normalized to GAPDH/Tubulin for other qPCR, and normalized to respective RNA amounts for cellular fractionation products, and $2^{-\Delta\Delta C_t}$ was used to calculate the fold expression.

5.5 Immunoblot

Following transfection, cells were harvested with trypsin and washed once with 1x PBS. Cells were then mixed with cell lysate buffer, protein inhibitor cocktail (539134, Millipore Sigma), sodium fluoride, and incubated in 95°C for 10 min before loading into a protein gel. Cell lysates were loaded into a 10-well 10% Bis-Tris SDS-polyacrylamide gel electrophoresis (Invitrogen, USA) along with PageRuler Plus Protein ladder (26619, ThermoFisher Scientific). Proteins were separated at 130V for 90 mins and subsequently transferred into a nitrocellulose membrane using standard SDS-PAGE protocols. After, nitrocellulose membranes were blocked with the LI-COR Biosciences Intercept locking buffer (VWR International) for 1hr at room temperature and incubated with primary antibodies overnight at 4°C on a rocker. The following primary antibodies were used: p65 (10745-1-AP, Proteintech), p-p65(3033T, Cell Signaling Technology), STAT5(25656T, Cell Signaling Technology), p-STAT5(80115-1-RR Proteintech), NKAP (ab229096, Abcam), GAPDH (60004-1-Ig, Proteintech), p53(10442-1-AP, Proteintech), Actin(66009-1-Ig, Protechtech) , Histone-3 (4499S, Cell Signaling Technology), and γ H2AX (60566SF, Cell Signaling Technology) . All antibodies were diluted using 1x TBST. After incubation, membranes were washed four times for 5 minutes with 1x TBST (0.1% Tween 20 TBS) buffer, The nitrocellulose membrane was incubated with goat Anti-mouse (102971-154, VWR International) and goat Anti-rabbit (103011-498, VWR International) IRDye-conjugated secondary antibodies for 1 hour under dark conditions on a rocker. After four 5-minute washes,

the membrane was analyzed via Odyssey Fc device (LI-COR Bioscience, USA) and analyzed with Empiria Studio software (LI-COR Bioscience, USA) on channels 700 and 800 for 2 mins each.

5.6 Cellular Fractionation

1×10^7 A549 cells were collected by trypsinization and resuspended in 1mL hypotonic lysis buffer (10 mM Tris-HCl [pH 7.5], 10 mM NaCl, 3 mM $MgCl_2$, 0.3% NP-40 supplemented with 1 complete protease inhibitor cocktail (Roche) and 10% glycerol). After 10 minutes of incubation on ice and a brief vortex of the sample, 60 μ L input control for Western Blot was saved and the rest was centrifuged for 8 minutes, 800 rcf at 4°C. The supernatant was saved as the cytoplasmic fraction, with 5M NaCl added to equal 140mM concentration. The remaining pellet was washed by resuspending in the hypotonic lysis buffer and centrifuged for 2 minutes at 300 rcf for 4 times, the supernatants were saved as washes. The total nuclear extract was then prepared by suspending the pellet in 0.5mL nuclear lysis buffer (20 mM Tris-HCl [pH 7.5], 150 mM KCl, 3 mM $MgCl_2$, 0.3% NP-40 supplemented with 1 complete protease inhibitor cocktail). The input sample was bath-sonicated for 15 minutes while the nuclear fraction and the four washes were sonicated with microtip at 25% duty and 4 output for 60 seconds each. To prevent the samples from warming during the process, they were placed in a beaker filled with ice. After sonication, all samples except input were put into the centrifuge at 18000 rcf for 10 minutes at 4°C and the supernatants were saved into new tubes and saved as samples for Western Blot and RNA isolation. U1 and GAPDH were probed as reference for nuclear and cytoplasmic markers, respectively.

5.7 Colony Formation assay

0.2M A549 cells were seeded at a 6-well plate overnight, followed by LNA GapmeRs knockdown targeting GRAS1. After 24 h incubation in a cell culture incubator at 37 C and 5% CO₂, transfected cells were collected using trypsin (Corning, USA) and resuspended using 4 ml EMEM medium. 50 ul of resuspended cell solution was added to 36.5 cm² cell dishes with 4 ml EMEM medium and was grown at the cell culture incubator for 10 days. Then, the medium was discarded and the cells were fixed using a 500 ul fixation buffer (methanol: acetic acid = 3 : 1). Then the fixed cells were stained by 0.1% crystal violet (Ricca Chemical, TX, USA).

5.8 Immunofluorescence

0.1M cells were seeded in a tissue culture dish with glass bottom (20102018, FluoroDish, FL, USA) overnight before transfection. After transfection, cells were washed three times with ice-cold PBS. Cells were fixed with chilled 100% methanol at room temperature for 5 minutes, then washed three times with PBS for 5 minutes. Cells were permeabilized with 0.2% Triton X-100 for 10 minutes at room temperature, then washed three times with PBS for 5 minutes. Then, cells were incubated in 1% BSA (bovine serum albumin) in PBST (PBS + 0.1% Tween 20) for 30 minutes blocking using Normal Goat Serum (ab7481, Abcam) at room temperature. Cells were incubated in primary antibodies for 1 hour at room temperature and washed three times with PBS. Then cells were incubated with Goat Anti-Rabbit IgG H&L Secondary Antibody (ab6717, Abcam) for 1 hour and washed 3 times with PBS. Finally, cells were incubated in either 1 µg/mL DAPI DNA stain (4083S, Cell Signaling Technology) or 0.5 ug/ml Hoechst 33342 Reagent (R37165, ThermoFisher Scientific, USA) for 1 min, rinsed with PBS, and imaged using the Eclipse Ti Inverted Microscope (Nikon, Japan) with Nikon's NIS-Elements imaging software.

5.9 Protein Immunoprecipitation

For each reaction, 3.5×10^7 A549 cells were harvested and resuspended in 1X PBS, with equal amount of cold polysome lysis buffer added (20 mM Tris-HCl [pH 7.5], 50 mM KCl, 10 mM MgCl_2 , 1 mM DTT, 1 complete protease inhibitor cocktail, 200 units/mL RNase inhibitor murine (New England Biolabs)). Cells were lysed by incubation at 4 °C for 5 minutes followed by incubation at -80 °C for about two hours before centrifuging for 10 minutes at 20000 rcf. The supernatants were collected and divided so that there were 1×10^7 A549 cells per sample and 5×10^6 cells saved as whole cell lysate inputs. 75 μL /sample of Dynabeads Protein G (Thermo Fisher Scientific) were washed twice with NT-2 buffer (50 mM Tris-HCl [pH7.5], 150 mM NaCl, 1mM MgCl_2 , 0.05% NP-40) and resuspended in 100 μL /sample of NT-2 buffer, with 3 μg of each antibody (IgG as a negative control (10284-1-AP, Protechtech), TDP43 as a positive control (10782-2-AP, Proteintech), NKAP, and STAT5) used. The beads/antibody mixtures were incubated for 90 minutes at room temperature, and then the beads were washed six times with NT-2 buffer before resuspending in NET-2 buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1mM MgCl_2 , 0.05% NP-40, 20 mM EDTA [pH 8.0], 1mM DTT, 200 units/mL RNase inhibitor murine). Each antibody had the cell lysate supernatant added so the total volume including the NET-2 would be 1 mL. All 3 immune complexes were incubated overnight at 4 °C with gentle rotation, followed by six washes with NT-2 buffer, then split in half for protein and RNA analysis. For all RNA samples, the beads were resuspended in NLS elution buffer (20 mM Tris-HCl [pH 8.0], 10 mM EDTA [pH 8.0], 2% NLS, and 2.5 mM TCEP) for proteinase K digest, then the beads were separated and standard RNA silane cleanup with Dynabeads Myone Silane (Thermo Fisher Scientific) was performed on the supernatant before reverse transcribing the

RNA using SuperScript III to perform qPCR. For the protein samples, 10 μ L of 2 Tris-Glycine SDS sample buffer (Thermo Fisher Scientific) was used to resuspend the beads and incubated at 50 °C for 10 minutes, then the elution was added to 9 μ L of ultrapure water and 1 μ L of 0.1 M DTT, the 20 μ L was saved as elution 1. Elution 2 was collected by resuspending the beads in 10 μ L of 2 Tris-Glycine SDS sample buffer and 1 μ L of 0.1 M DTT and incubating at room temperature for 5 minutes before collecting the elution and adding 9 μ L of ultrapure water. 15 μ L of the input was also prepared for Western blot by adding 2 Tris-Glycine SDS sample buffers to 1 concentration. All protein samples were heated at 95 °C for 10 minutes before storing at -80 °C.

5.10 MTT cell survival assay

the cell viability on A549, HCT116, MCF7, K562 and HEK293 cells was determined by modified microculture tetrazolium (MTT) assay (University at Buffalo, State University of New York, NY, USA), based on the enzymatic reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide to form formazan crystal by mitochondria and cellular dehydrogenase enzymes. Before any treatment, 15k cells were seeded each well into 96-well flat-bottomed microplates containing medium overnight. The cells were treated/transfected and incubated in a humidified incubator in 5% CO₂ for 48h at 37°C, and then replaced the medium with the working MTT solution following the manufacturer's instructions; the resulting solution was incubated in 5% CO₂ for another 20 mins at 37°C. Formazan crystals were dissolved in 100 μ L of DMSO. The plates were then analyzed on the Synergy HTX Multi-Mode Microplate Reader (Agilent, USA) at 570 nm. Each sample point was assayed with biological triplicate per assay, and cell line experiments were performed at least three times.

5.11 Cell cycle analysis

Following treatment, A549 cells were harvested with trypsin and washed twice with PBS. Cells were then permeabilized and fixed with 70% ice-cold ethanol drop-wise and incubated for 15 minutes on ice, gently vortexing to prevent clumping. Cells were washed twice with PBS. Then, cells were resuspended in PI Staining Solution (3.8mM sodium citrate, 50 µg/ml PI (Biotium, USA), 40 µg/ml RNase A (Zymo Research, USA)) with RNase and incubated for 40 mins at 37 °C in the dark. Cells were centrifuged and the supernatant was gently aspirated. Finally, the cells were resuspended in 500 ul ice cold PBS and examined on the S3e Cell Sorter (Bio-Rad Laboratories). The data of cell cycle profiles was analyzed by FlowJo software (BD Biosciences).

5.12 RNA sequencing

RNA from linc01977 overexpression at MCF7, linc01977 knockdown at MCF7, linc00883 overexpression at K562, PDCD4-AS1 knockdown at K562, GRAS1 overexpression at A549, and GRAS1 knockdown at A549 with each control samples respectively was isolated by RNeasy Mini kit and purified by RNA clean & concentrator kit. Illumina RNA-seq libraries were done by the Institution for Genomic Medicine (IGM) at UCSD using Illumina stranded mRNA prep. Quality of the libraries was examined, and the sequencing is done in the NovaSeq S4 with the run type of PE100.

5.13 RNAseq analysis

Raw paired end fastq sequencing reads were mapped with STAR (version 2.7.9a) and a genome generated with STAR's genomeGenerate to the human reference genome downloaded from UCSC genome browser (GRCh38) and human gene annotation file downloaded from genecode (GRCh38, version 38 (Ensembl 104)). The output bam files are run through FeatureCounts (version 2.0.3, from Subreads) to generate a matrix of read counts. This matrix is put into DESeq2 (version 1.34.0) to get differential gene expression data. Further gene ontology (GO) enrichment analysis was done with enrichGO (from clusterProfiler version 4.2.2). Relevant ENSG were queried with getBM (from biomaRt library) to obtain the HGNC gene symbols.

5.14 RNA antisense purification - mass spectrometry (RAP-MS)

RAP on whole cell lysis: Each RAP-MS experiment used 300M cells (200M per GRAS1, 50M per U1 as a positive control, 50M per luciferase as a negative control). Cells were collected and crosslinked with 254 nm wavelengths by XL-1500 Spectrolinker (Spectro-UV, USA). Per 20 M cells, the cell pellet was resuspended with 900 ul of cold lysis buffer (10 mM Tris-HCl pH7.5, 500 mM LiCl, 0.5% n-dodecyl -D-maltoside, 0.2% sodium dodecyl sulfate, 0.1% sodium deoxycholate) with 1X protease inhibitor cocktail and 920 U of murine RNase inhibitor (M0314L, New England Biolabs, UK) to prevent proteins and RNAs from degradation. Resuspended cell pellet was further lyzed in ice by passing through needles 10 times and 2 sets of 1.5 mins sonication (setting: 35 duty, 3.0 power output). Then the cell lysis was incubated 20 U TurboDNase (Invitrogen, USA) for 15 mins at 37 C to remove the DNAs in the lysis, and immediately quenched the DNase reaction by adding 10 mM EDTA, 5 mM EGTA, and 2.5 mM TCEP. 6 M Urea Hyb Buffer was added to the lysis to a final concentration of 4 M for denaturing the proteins. Afterwards, the lysis of 20 M cells was pre-cleared by washing with 125

ul of C1 Streptavidin beads (Invitrogen, USA) at 45 C for 30 mins. Biotinylated 90-mer DNA probes for the RNA (10 ug of probes for GRAS1, 5 ug of probes each for U1 and Luciferase) was mixed with pre-cleared lysate and incubated at 45 C for 2 hrs with mixing at 1000 rpm (15 s shaking, 15 s off). Complexes of biotinylated DNA probes and RNA were then bound to C1 Streptavidin beads by incubating at 45 C for 1 hr with the same mixing setting as above. Beads then were washed 4 times with 4 M Urea Hyb Buffer, and resuspended with 0.5 ul of Benzonase (E1014, SIGMA-ALDRICH) in 500 ul Benzonase Elution Buffer (20 mM Tris pH 8.0, 2 mM MgCl₂, 0.05% NLS, 0.5 mM TCEP) to release the proteins. Proteins were then precipitated by adding 80% TCA to final concentration to 20% for overnight incubation at 4 C followed by a 15-minute centrifuge with 10000 x g at 4 C.

Mass spectrometry preparation: The protein pellet was resuspended with 50 ul of water with 5% Rapigest (186002123, Waters) and incubated at 100 C for 10 mins. 0.5 ug of trypsin was added with 37 C overnight incubation with 600 rpm shaking for protein digestion. Protein samples were cleaned up by HiPPR detergent removal resin (PI88305, VWR International) and dried completely by speedvac (Savant) before mass spectrometry analysis at Sanford Burnham Prebys Proteomics Core.

5.15 Annexin V apoptosis assay

Following treatment or transfection, A549 cells were collected via trypsin and washed in cold PBS. Then cells were resuspended in 100 ul of 1X Annexin-binding buffer (10 mM Hepes adjusted to pH 7.4, 140 mM NaCl and 2.5 mM CaCl₂) with 5 ul of FITC-conjugated Annexin V (A13199, Life Technologies) and 2ul of 1mg/ml Propidium Iodide (40017, Biotium). The cell solution was incubated for 30 mins at room temperature in the dark. After the incubation, 400 ul

of 1X Annexin-binding buffer was added to the cell solution and analyzed by the S3e Cell Sorter. The raw data was analyzed by FlowJo software.

5.16 Ferroptosis analysis

Following treatment or transfection, 1ul of 3 mM BODIPY-C11 581/591 (D3861, Life Technologies) was added to each well in the 6-well plate, and returned the plate to the tissue culture incubator for 3 hrs. After the incubation, cell media with BODIPY added was transferred to a 15 ml tube, and the cells was washed with 1 ml HBSS buffer (140 mM NaCl, 5 mM KCl, 1 mM CaCl₂, 0.4 mM MgSO₄, 0.5 mM MgCl₂, 0.3 mM Na₂HPO₄, 0.4 mM KH₂PO₄, 6 mM D-Glucose, and 4 mM NaHCO₃). The washed HBSS and cells harvested by trypsin were collected in the corresponding 15 ml tube and centrifuged the tube at 0.3 rcf for 3 mins. The supernatant was removed, and the pellet was resuspended with 500 ul HBSS for FACS analysis using the S3e Cell Sorter. The raw data was analyzed by FlowJo software.

Table 5.1 List of antibodies

Antibodies	Source	Identifier
Lamin A/C Polyclonal antibody	Proteintech	Cat No. 10298-1-AP
GAPDH Monoclonal antibody	Proteintech	Cat No. 60004-1-Ig
P53 Polyclonal antibody	Proteintech	Cat No. 10442-1-AP
NF-κB p65 Polyclonal antibody	Proteintech	Cat No. 10745-1-AP
Cytochrome c Polyclonal antibody	Proteintech	Cat No. 10993-1-AP
Ubiquitin (P4D1) Mouse mAb	Cell Signaling Technology	Cat No.#3936
Phospho-Histone H2A.X (Ser139) (20E3) Rabbit mAb	Cell Signaling Technology	Cat No.#9718
Anti-NKAP antibody	Abcam	Cat No. Ab229096

Table 5.2 List of qPCR primers

RT-qPCR primers		
Target	Forward (5'-3')	Reverse (5'-3')
hGAPDH	GGG CTC TCC AGA ACA TCA TCC	GTG CAC CAC TGA CAC GTT GG
TUBB	GAT CAG TGA TGA ACA TGG CAT CG	AAA GGA CCT GAG CGA ACA GAG
TUBB	CCA GAC AAC TTT GTA TTT GGT CAG T	CGT ACC ACA TCC AGG ACA GAA T
PML	CCG TCA TAG GAA GTG AGG GCT TC	GTT TTC GGC ATC TGA GTC TTC CG
NKAP	GCT GGA AAA CGT ATC CCA CGA AG	CTG GTT CTC TTT TCG CAG TCG C
BDKRB1	GCC TGT GAC AAT GCT CCA GAA G	TTT CTG CCA CGT TCA GTT GCC G
PLK3	CTA TGT GGC TCC AGA AGT GCT G	TCC TTC AGG TCA GCC GTC TCA A
GADD45a	TGC TCA GCA AAGCCC TGA GT	GCA GGC ACA ACA CCA CGT TA
GCLC	GGA AGT GGA TGT GGA CAC CAG A	GCT TGT AGT CAG GAT GGT TTG CG
COX2	CGG TGA AAC TCT GGC TAG ACA G	GCA AAC CGT AGA TGC TCA GGG A
BACE1	GTG AGG TTA CCA ACC AGT CCT TC	CGT GGA TGA CTG TGA GAT GGC A
ANGPT1	CAA CAG TGT CCT TCA GAA GCA GC	CCA GCT TGA TAT ACA TCT GCA CAG
NR3C1	GGA ATA GGT GCC AAG GAT CTG G	GCT TAC ATC TGG TCT CAT GCT GG
hU1	TTA CCT GGC AGG GGA GAT AC	TCC CAC ATT TGG GGA AAT C
GRAS1	CCT TTT GGT GCG AGC TTG CTG T	GTC GAC TCA CCA GAC TCA GGG TG

Table 5.3 List of cell lines

Experimental model: Cell line		
Human cell line	Source	Identifier
A549	ATCC	No. CCL-185
K562	ATCC	No. CRL-3344
MCF7	ATCC	No. HTB-22
HEK293	ATCC	No. CRL-1573