UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Exploration of NKAP as a lncRNA-Binding Protein

Permalink

https://escholarship.org/uc/item/05q945q7

Author

Camden, Nichelle Sierra

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Exploration of NKAP as a lncRNA-Binding Protein

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Chemistry

by

Nichelle Sierra Camden

Committee in charge:

Professor Colleen McHugh, Chair
Professor Elizabeth Komives
Professor Alexis Komor

2024

Copyright

Nichelle Sierra Camden, 2024

All rights reserved.

The Thesis of Nichelle Sierra Camden is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

Although they can't read, I would like to dedicate my thesis to my cats, Kaiju and Stormy, for their irreplaceable support. I would have never moved across the country to attend UCSD without them. I wouldn't be here without their unconditional love, constant care, and silly little mannerisms.

I would also like to dedicate my thesis to a few notable teachers, although I've been fortunate to have many: to my high school AP biology teacher, Mr. Bichler, for inspiring my curiosity and passion for mechanisms in biology and biochemistry; and to my undergraduate professors and mentors (for many of our encouraging day-to-day interactions, but especially), Dr. Michael Wentzel, for persistently encouraging me to apply to this program, and to Dr. Rebekah Dupont, for helping me get through the pandemic lockdown with as much sanity intact as anyone could have hoped for. Your influences, compassion, and mentorship have been invaluable boons in my life and career, and I am forever inspired to have a similar impact on others.

TABLE OF CONTENTS

THESIS APPROVAL PAGE iii

DEDICATION iv

TABLE OF CONTENTS v

LIST OF FIGURES vi

LIST OF TABLES viii

LIST OF ABBREVIATIONS ix

ACKNOWLEDGEMENTS x

VITA xi

ABSTRACT OF THE THESIS xii

Chapter 1 Introduction 1

Chapter 2 Biological importance of STAT5 and pivot to NKAP 4

Introduction 4

Methods 5

Results 10

Conclusions 17

Acknowledgements 17

Chapter 3 NKAP Functional Assays 18

Introduction 18

Methods 18

Results 19

References 42

LIST OF FIGURES

Figure 1.1 Central dogma updated to include ncRNA. 1

Figure 1.2 New analysis reveals the abundance of lncRNAs. 2

Figure 2.1: RNA-antisense purification-mass spectrometry (RAP-MS) of the GRAS1 lncRNA in cancer cell lines. 5

Figure 2.2 General schematic showing the steps of an immunoprecipitation experiment. 11

Figure 2.3 Western blot showing the detection of STAT5 and TDP-43 following an immunoprecipitation. 12

Figure 2.4 Bar graph showing the fold change expression of TARDBP, BARD1, and GRAS1 following an immunoprecipitation of STAT5, TDP-43, and IgG. 12

Figure 2.5 General scematic of iCLIP data analysis. 14

Figure 2.6 Sequencing reads from the Lin Lab iCLIP align to FTH1 and FTL. 14

Figure 2.7 Amplification plots of NKAP targets showing effective qPCR primers. 15

Figure 2.8 Boxplot of GRAS1 gene expression after transfecting with GRAS1 siRNA. 16

Figure 2.9 GRAS1 gene expression knockdown in A549 cells measured by qPCR. 16

Figure 3.1 Schematic diagram of plasmid construction for NKAP expression. 20

Figure 3.2 Results of PCR amplification of NKAP from cDNA 21

Figure 3.3 DNA gel and plasmid map of overexpression construct CMV-NKAP. 22

Figure 3.4 Restriction enzyme digest screening results showing successful insert of NKAP fragment. 23

Figure 3.5 PCR amplification of 3x Myc tag and optimization of NKAP PCR amplification. 24

Figure 3.6 DNA gel and plasmid map of overexpression construct CMV-Myc-NKAP. 25

Figure 3.7 DNA gel showing the PCR amplification of NKAP and EGFP for construction of NKAP-EGFP fusion and EGFP control expression plasmid. 26

Figure 3.8 Plasmid map and verification of NKAP overexpression contruct and EGFP control. 27

Figure 3.9 Microscope images of A549 cells taken in the blue channel after transfecting with CMV-EGFP, CMV-EGFP-NKAP, or CMV-NKAP. 29

Figure 3.10 Western blot of A549 cell lysate following a transfection, probed for EGFP. 30

Figure 3.11 Transfection of NKAP overexpression plasmids increases NKAP mRNA. 31

Figure 3.12 Microscope images of A549 cells transfected with the expression vectors CMV-ABE-P2A-EGFP, pcEGFP, pMAX-GFP, CMV-EGFP-NKAP, or CMV-EGFP and Lipofectamine2000 33

Figure 3.13 Two plasmid maps highlighting the addition of the Kozak sequence in the design for CMV-Myc-NKAP-P2A-EGFP and CMV-EGFP-NKAP. 34

Figure 3.14 Successful PCR amplification of NKAP, Myc-NKAP, and P2A plasmid backbones. 35

Figure 3.15 DNA gel screening plasmids for the correct size before sequencing. 36
Figure 3.16 Plasmid maps of CMV-EGFP-NKAP-bGH and CMV-Myc-NKAP-P2A-EGFP-bGH 36
Figure 3.17 Comparison of EGFP expression in plasmids. 37
Figure 3.18 Western blot of A549 cell lysate following a transfection of CMV-Myc-NKAP-P2A-EGFP, CMV-EGFP-NKAP, CMV-Myc-HDM2, or CMV-ABE-P2A-EGFP. 38
Figure 3.19 Western blot of A549 cell lysate following a transfection of CMV-EGFP-NKAP, CMV-Myc-NKAP-P2A-EGFP, CMV-Myc-HDM2, or CMV-ABE-P2A-EGFP..... 39
Figure 3.20 Final NKAP overexpression constructs. 40

LIST OF TABLES

Table 1 Optimizing transfecting EGFP-containing constructs into A549 cells in a 6-well plate. . 28
Table 2 NKAP gene fold change expression values of CMV-EGFP-NKAP measured by qPCR. 31
Table 3 Summary of the overexpression plasmids I created and their important features. 39

LIST OF ABBREVIATIONS

lncRNA	long non-coding RNA
NKAP	NF-kB Activating Protein

ACKNOWLEDGEMENTS

I would like to thank Professor Colleen McHugh for giving me the opportunity to learn about RNA biochemistry in her lab, and for all the advice and support she offered throughout my Master's thesis. Also thank you to the rest of the McHugh lab, especially Vita, Simone, Jonathan, Adarsh, and Lauren for their help and advice.

I would like to thank the ACS Bridge Program for supporting my thesis, especially Professor Brian Leigh for all his work in ensuring the ACS Bridge can continue its mission for future cohorts, and for cementing in my mind "don't sweat the small things". I would also like to thank Professor Elizabeth Komives for all of her time and energy in helping me organize my data and draft my thesis, and for her continual encouragement and advice. I would not have found such a wonderful home at UCSD without their support.

I would like to acknowledge Dr. Tong Su and Jonathan Zhu for their contributions in chapters 1 and 2. I would also like to thank Jonathan for his patient and thorough explanations of his bioinformatics analyses. I'm also very grateful to Vita Chou for being so wonderful to work with. I've learned so much throughout our conversations and it's really been an honor to be Vita's mentor.

Finally, I would like to thank my friends and family, especially my mom and grandma for always being such strong advocates for education. I also want to thank Dr. Quinn Cowan, for being the very best partner. I'm so lucky to have such insightful and encouraging people in my life.

Chapter 2, in part, is a reprint of the material as it appears in bioRxiv . Su, Tong; Trang, Nu; Zhu, Jonathan; Kong, Lingbo; Cheung, Darin; Chou, Vita; Ellis, Lauren; Camden, Nichelle; Huang, Calvin; McHugh, Colleen. GRAS1 non-coding RNA protects against DNA damage and cell death by binding and stabilizing NKAP. bioRxiv, 2024. The thesis author was a contributing investigator and author of this paper.

VITA

2021 Bachelor of Science in Biochemistry, Augsburg University

2024 Master of Science in Chemistry, University of California San Diego

FIELD OF STUDY

Major Field: Chemistry
Studies in lncRNA Biochemistry
Professor Colleen McHugh

ABSTRACT OF THE THESIS

Exploration of NKAP as a lncRNA-Binding Protein

by

Nichelle Sierra Camden

Master of Science in Chemistry

University of California San Diego, 2024

Professor Colleen McHugh, Chair

I explored the interactions of Nf-kappa B activating protein and the noncoding RNA GRAS1. I knocked down GRAS1 and studied the effects on NKAP target genes using qPCR. I also overexpressed NKAP in A549 cells using overexpression plasmids, including a Myc-tagged, P2A-EGFP construct, and an EGFP-fusion construct.

Chapter 1 Introduction

According to the Central Dogma, genetic information is transcribed from DNA to RNA, then translated into proteins. However, only about 2% of the human genome encodes for proteins, while noncoding RNA is estimated to make up about 80% of the human genome[1]. For many decades, noncoding genes were considered "junk DNA" whose functions were unknown, but thanks to efforts like the Human Genome Project and ENCODE, there have been great advancements in understanding the abundance and functions of noncoding RNAs[2].

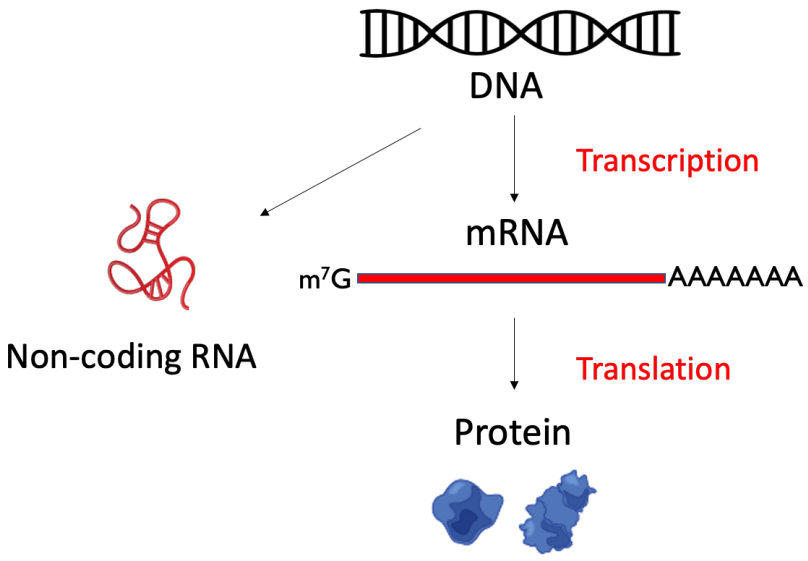


Figure 1.1 Central dogma updated to include ncRNA.

Noncoding RNAs are very abundant, consist of different categories, and hold many functions. Some examples of the different types of ncRNA include long noncoding RNA (lncRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), small interfering RNA (siRNA), as well as many others. lncRNA is somewhat arbitrarily defined as noncoding RNA genes whose transcripts are 200 nucleotides or longer[3]. Due to early RNA purification methods, 200 nucleotides was a convenient size cut-off for categorizing lncRNAs as it was long enough to exclude many of the

shorter RNA transcripts such as tRNA and 5S rRNA[3]. Colloquially, this classification mainly works, but oversimplifies and obscures new insights in ncRNA. Recently, it's been suggested to reclassify lncRNAs as noncoding transcripts 500 nucleotides or longer that are mostly transcribed by RNA polymerase II (Pol II), which also transcribes all mammalian protein-coding genes[4]. Many lncRNAs, like proteins, undergo splicing, polyadenylation, and are capped with 7-methylguanosine when transcribed by Pol II[5]; although some, like 18S and 28S rRNAs, are transcribed by Pol I, and others by Pol III[6].

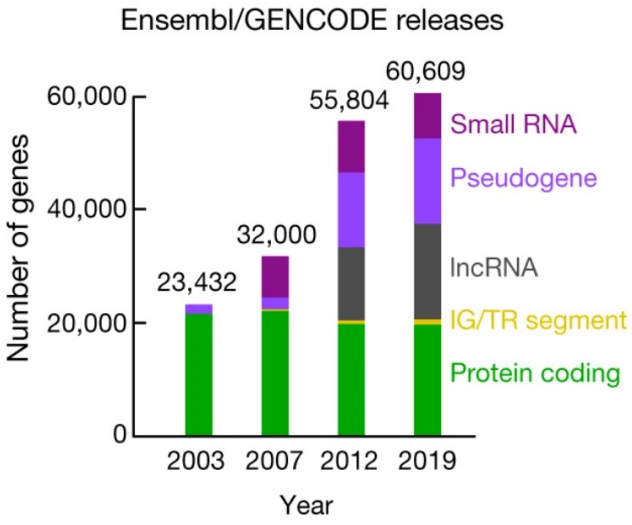


Figure 1.2 New analysis reveals the abundance of lncRNAs.[2] As genome mapping projects have advanced, so has our understanding of the abundance of lncRNAs.

Some functions of lncRNAs include acting as a molecular decoy, scaffold, guide, and signal, and being involved in epigenetic and post-transcriptional gene regulation, as well as regulating chromatin states[7]. Because lncRNAs play such an important role in modulating and regulating various pathways, when they are misregulated or expressed aberrantly, the physiological impacts can be very consequential.

LncRNAs play critical roles in many diseases, including various types of cancer and neurological diseases. There have been lncRNAs confirmed to be involved in important hallmarks of cancer, contributing to both the onset and progression of cancer.[8] Our lab is investigating the

lncRNA Growth Regulator Antisense 1 (GRAS1), which has been found to play a role in the viability of cancer cells. [9]

GRAS1 is expressed in many cell types from the EPB41L4A-AS1 locus, located on Chromosome 5q22.1, and has been found in both the nuclear and cytoplasmic fractions of cells.[9] Our lab's finding that GRAS1 plays a role in cell survival and viability is further supported by analysis of RNA expression using the Cancer Genome PanCancer Atlas, which shows that GRAS1 expression is often dysregulated in multiple types of cancer.[10,11]

Another important contributor to diseases that we are investigating is Nf-kappa B activating protein (NKAP). Located on the X Chromosome, NKAP is ubiquitously expressed and involved in the activation of NF-κB.[12] A gene ontology of NKAP shows RNA and chromatin DNA-binding, specifically acting as a transcriptional corepressor of Notch target genes, and is associated with the histone deacetylase HDAC3.[13] NKAP has also been found to promote cancer growth in colon, pancreatic, and glioma cancers, even protecting glioblastoma from cell death via ferroptosis.[14–18] Additionally, NKAP has also been found to contribute to various developmental and cognitive impairments, including Marfanoid habitus, scoliosis, and Hackmann-Di Donato-type X-linked syndrome, as well as congenital heart disease.[19,20]

Chapter 2 Biological importance of STAT5 and pivot to NKAP

Introduction

After analyzing the results of a CRISPR knockdown screen[21] and assessing which genes were differentially expressed in cancer cells, a senior graduate student in my lab, Tong Su, created a list of lncRNAs suspected to be involved in regulating cancer growth in many different tissue types. To investigate which lncRNAs contributed to a growth phenotype in several immortalized mammalian cell lines, all originating from cancer, he designed experiments to overexpress and knockdown each lncRNA. Testing each lncRNA separately, he cloned a lncRNA gene into an overexpression vector and designed a GapmeR to transfect into mammalian cells, measuring the respective overexpression and knockdown efficiencies of each target via qPCR. In this context, a GapmeR is an oligonucleotide that is designed to be antisense to a target endogenous RNA, binding through base complementarity, and eliciting degradation by recruiting ribonuclease H.[22] He decided an overexpression plasmid was efficient if it increased the lncRNA by 3 fold or greater, and a knockdown experiment was efficient if it reduced the target lncRNA by 10% or more. My project is a continuation of Tong's work to investigate the lncRNA EPB41L4A-AS, also known as Growth Regulator Antisense 1 (GRAS1).

Tong showed via RAP-MS that GRAS1 RNA interacts with STAT5 in K562 cells. Figure 1 shows the top proteins identified in a RAP-MS capture of GRAS1 in K562 and A549 cells, respectively. RAP-MS is a method developed by our PI, Colleen McHugh, that allows uncharacterized RNA molecules to be isolated so that the proteins they bind to can be determined using Mass Spectrometry, giving clues to their functions.[23]

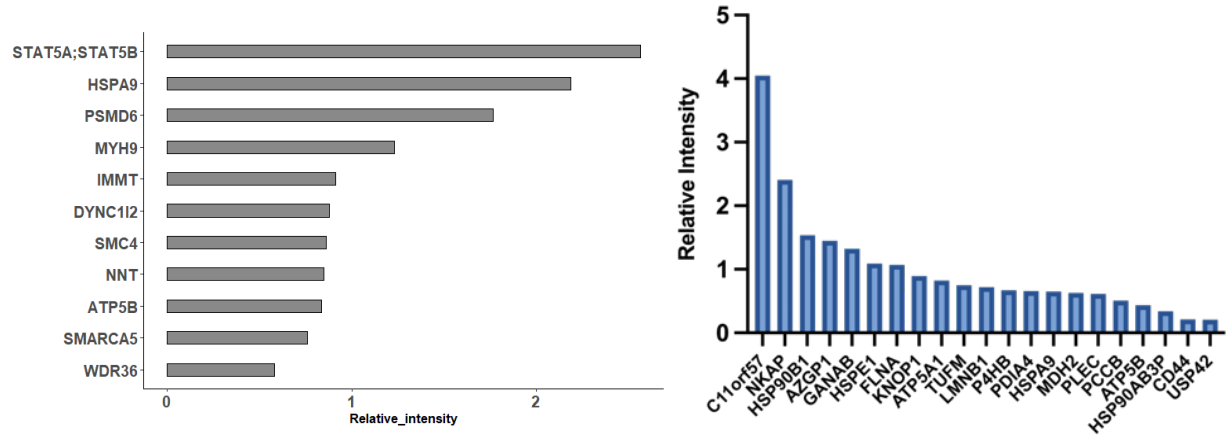


Figure 2.1: RNA-antisense purification-mass spectrometry (RAP-MS) of the GRAS1 lncRNA in cancer cell lines. (Left) Top proteins identified in a RAP-MS capture of GRAS1 in K562 cells. (Right) Top proteins identified in a RAP-MS capture of GRAS1 in A549 cells.

Methods

Immunoprecipitation. Preparation of Whole Cell Lysate

For each sample, I prepared 1.1 x10^7 A549 cells by washing them with 1x phosphate-buffered saline (PBS) and lysing them with polysome lysis buffer (PLB). The PLB was composed of 1mM 0.1M DTT, 200 units/mL RNAse inhibitor murine cocktail (New England Biolabs), 1x complete protease inhibitor cocktail (Roche), and ultra-pure (UP) water (Invitrogen). I incubated the cells and PLB at 4°C for 5 minutes, then at -80°C for two hours. I thawed the whole cell lysate (WCL) samples on ice before centrifuging them for 10 minutes at 20,000x g in at 4°C. I collected the supernatants and set aside 5% of the WCL to use as an input later.

Preparation of Magnetic Beads & Immobilization of Antibodies

I used 75 µL of Dynabeads Protein G (Thermo Fisher Scientific) per sample. I washed the beads twice with 500 µL of NT-2 buffer, and then resuspended them in 100 µL per sample of NT-2 buffer. The NT-2 buffer was composed of 50 mM Tris-HCl at pH7.5, 150 mM NaCl, 1 mM $MgCl_2$, and 0.05% NP-40. I added 5 µg of the corresponding antibody to each sample: STAT5 (ab264324, Abcam), TDP-43 as a positive control (10782-2-AP, Proteintech), and IgG as a negative control (30000-0-AP, Proteintech). I incubated the beads and antibody mixtures at room temperature for

1.5 hours with rotation (orbital 15/15, reciprocal 5/11, no vibration). After incubation, I centrifuged each sample for 15 seconds at 5000x g, and placed them on a magnetic rack for 1 minute. I removed the supernatants (antibody flow-throughs), and stored at -80°C for later. I washed each bead-antibody sample a total of 6 times with ice-cold NT-2 buffer.

I prepared NET-2 buffer, which was composed of 1x NT-2 buffer, 0.5M EDTA (pH 8.0), 1 mM DTT, 200 units/ mL RNAse inhibitor, and UP water. I calculated what volume of my WCL supernatant would be equivalent to 10 million cells, then added the appropriate volume of NET-2 buffer to a final volume of 1 mL. I resuspended each of my washed beads-antibody samples with 1 mL of the supernatant-NET-2 solution. I incubated each sample tube with rotation using the same settings previously described overnight at 4°C.

Washing the Protein-RNA Complex on Beads

The following day, I briefly centrifuged the samples to spin down the contents, and placed the tubes on a magnetic rack in ice for 1 minute. I removed and saved the supernatant (complex flow-through) and stored at -20°C to use later. I added 1 mL of ice-cold NT-2 buffer to each sample and vortexed to resuspend the beads. After a brief spin in the microcentrifuge, I placed the tubes on the magnetic rack on ice for 1 minute. I removed and saved the first wash, storing at -20°C. This wash step was repeated 5 more times for a total of 6 washes, after which I discarded the remaining washes. Before removing the final wash, I split the samples equally into two aliquots each (2 x 500 µL each) for protein and RNA analysis.

RNA Sample Clean-up: Proteinase K Digest

For the RNA samples, I retrieved the WCL supernatant input from -80°C and thawed it on ice. I transferred 5 µL (equivalent to 5,000 cells) into an Eppendorf tube and added 15 µL of NLS pH 8.0 buffer. I removed the final wash and resuspended each sample in 18 µL of NLS pH 8.0 buffer before incubating the samples at 95°C for 2 minutes. The NLS buffer was composed of 20 mM Tris–HCl pH 8.0, 10 mM EDTA pH 8.0, 2% NLS, and 2.5 mM TCEP. After allowing the samples

to cool, I added 2 µL of Proteinase K and incubated them for 1 hour at 50°C with shaking at 1,100 rpm. The final reaction volume of each sample was 20 µL. I performed a standard RNA silane cleanup with Dynabeads MyOne Silane (Thermo Fisher Scientific), carefully allowing the beads to dry after the final wash, ensuring they became matte but not cracked. I resuspended each sample in 11 µL of UP water to elute the RNA into PCR tubes, recovering approximately 10 µL of RNA. The final RNA elutions were each reverse transcribed using SuperScript III so that I could perform qRT-PCR, and stored at -80°C.

Protein Elution

To elute the protein, I magnetically separated the beads from the other portion of the samples and removed the final wash. For 37.5 µL of beads (half of the initial volume), I resuspended each sample in 10 µL of 2X SDS buffer (without DTT) and incubated at 50°C for 10 minutes. In an Eppendorf tube, I prepared the elution buffer by adding 9 µL of UP water and 1 µL of 0.1M DTT. After incubation, I placed the tubes on a magnetic rack and transferred the elution from the beads to the Eppendorf tube (Elution 1). I repeated the elution process by adding 10 µL of 2X SDS buffer to the beads, then added 1 µL of 0.1M DTT. After placing the tubes on the magnetic rack, I added 9 µL of UP water to each sample and transferred the second elution from the beads (Elution 2). I denatured all the protein elution samples at 95°C for 10 minutes, and stored them at -80°C.

Cross-link immunoprecipitation (CLIP) data analysis method. Our pipeline to process CLIP datasets, created by lab member Jonathan Zhu, is shown in the following steps:

First, access the lab workstation computer remotely through command line. Most laptops are not suited for the level of computing needed to process CLIP datasets, so our workstation desktop computer is essential. Access the raw sequencing files and unzip them using the command gzip –d. Next, use the command fastqc –t 2 –o raw_qc/File_Name_1.fastq File_Name_2.fastq to check the quality of the dataset, and identify the sequencing adapters that

need to be trimmed off. To trim the adapters and bases with low scores, use the command cutadapt -j 24 -q 20 -m 20 -a and -o ../trim_fastq/File_Name_1.trim.fastq -p ../trim_fastq/ File_Name_2.trim.fastq File_Name_1.fastq File_Name_2.fastq. Also, remove trimmed sequences with a length less than 20 as they will be difficult to align. Save the trimmed fastq file in a different folder.

To align the reads to a reference genome, use the commands STAR --readFilesIn File_Name_1.trim.fastq File_Name_2.trim.fastq --outFileNamePrefix ../map/ File_Name_ --genomeLoad LoadAndKeep --genomeDir /data/programs/genome/star_ genome/[24]. To skip an additional step of converting the files to a BAM format, use the following command to output the aligned files in BAM: --outSAMtype BAM Unsorted --runThreadN 24. Next, use the samtools command to sort the alignment files to access them by genomic position instead of sequencing order. The following command is an example: samtools sort -@ 24 File_Name_Aligned.out.bam -o File_Name.sorted.bam[25]. To generate index files for the sorted alignment and quickly find positions on the genome, use the command: samtools index File_Name_Aligned.out.bam.

Write a csv file to tell the function lab_TPM.R to get TPM (transcripts per million) values for the newly indexed BAM file. Exit the csv file editor.

After aligning the reads to the reference genome, use the software HOMER to find the peaks. HOMER is recommended because it makes analyzing quicker and requires less memory by using a special type of processing that creates “tags” for the alignment files[26]. To start, use the command: makeTagDirectory ../peak_calling/tag_dir/ File_Name.sorted.bam. Then, to find the peaks, use the command findPeaks tag_dir/ -o File_Name_iclip_peaks.txt -style clip. Using the CLIP settings, HOMER will use the tag directory from the previous command to determine which regions should be called as peaks, although these thresholds can be adjusted. Finally, convert the file from the HOMER format into a BED file, which is viewable in Integrative Genomics Viewer (IGV) to visualize where the sequencing reads form peaks in the reference genome. To do this, use the command pos2bed.pl nk File_Name_iclip_peaks.txt > File_Name_iclip_peaks.bed.

Due to financial reasons, we decided to pivot from performing an eCLIP on STAT5 to investigating the other top proteins from the RAP-MS experiments with GRAS1 in A549 cells. We decided to explore the interaction between GRAS1 and Nf-kappa B activating protein (NKAP). Jonathan Zhu compared genes that were differentially expressed in RNA sequencing data of GRAS1 knockdown samples and genes found to be bound by NKAP in an iCLIP dataset produced by the Lin Lab[18]. Although the NKAP iCLIP experiment was done in pancreatic ductal adenocarcinoma (PDAC) cells, we hypothesized that it was a good starting point, and was the only high-quality NKAP CLIP dataset available at the time. Jonathan filtered between upregulated, downregulated, and all genes, and ran a gene ontology analysis of the three lists. He found that many of the genes bound by NKAP that were also differentially expressed when GRAS1 was knocked down were related to ferroptosis. Ferroptosis is a type of nonapoptotic cell death dependent on iron regulation[27]. From the NKAP iCLIP dataset, I created a short list of genes with high TPM values, confirmed they were also found in Jonathan's list of differentially expressed genes when GRAS1 was overexpressed or knocked down, and found that many of them played a role in iron-regulation, environmental stress response, or lipid peroxidation- all factors in ferroptosis. Figure 2.6 shows an example of reads from the Lin Lab NKAP iCLIP aligning to the genes Ferritin Heavy Chain 1 (FTH1) and Ferritin Light Chain (FTL), which code for proteins that make up ferritin, an essential cellular structure in iron storage and regulation. Using the peak calling pipeline outlined above, HOMER identified 1233 peaks from the Lin Lab NKAP iCLIP dataset.

Western Blot. I incubated the samples at 95°C for 10 minutes, then loaded 20 µL of each sample into a 10-well , 10% Bis-Tris SDS-polyacrylamide gel along with the PageRuler Plus Protein Ladder (26619, ThermoFisher Scientific) before separating the proteins by SDS-PAGE. Electrophoresis was carried out at 120V for 1 hour and 45 minutes, after which the proteins were transferred onto a 0.2 µm pore size nitrocellulose membrane.

The membrane was blocked with a 5% nonfat milk protein and TBS solution for one hour at room temperature. I cut the membrane according to what antibody I would use, and incubated the membranes overnight at 4°C on a rocker with the following primary antibodies: STAT5, GAPDH (60004-1-Ig, Proteintech), TDP-43, IgG, all diluted in 1X TBST.

The next day, I washed the membrane four times for 5 minutes each, with 1X TBST buffer (1X TBS with 0.1% Tween-20). The membranes were then incubated for one hour in the dark on a rocker with IRDye-conjugated secondary antibodies, including goat anti-mouse (102971-154, VWR International) and goat anti-rabbit (103011-498, VWR International). Following incubation, I washed the membranes four times for 5 minutes each with 1X TBST. Finally, I imaged the membranes using an Odyssey Fc device (LI-COR Biosciences, USA) and analyzed the results with Empiria Studio software (LI-COR Biosciences, USA).

Results

I wanted to optimize the IPs of STAT5 out of K562 cells in order to perform an enhanced cross-linked immunoprecipitation (eCLIP) experiment to identify bound RNAs. To do this, I first optimized the immunoprecipitation experiments. Protein immunoprecipitation (IP) is a method used to isolate a protein from cell lysate using magnetic beads coated in an antibody specific to the protein of interest. RNA-binding proteins can also be pulled down using this method to isolate the RNA molecules bound by a given protein. Radiating the sample with short-wave UV light can create covalent bonds between interacting protein and RNA, cross-linking and stabilizing the interaction, and providing information on binding sites when analyzed by high throughput sequencing.[28] The magnetic beads are mixed with the cell lysate, allowing the antibody to interact with the protein of interest, binding it to the magnetic beads as well. The protein-bead complexes are washed several times to remove nonspecific binding, and the bound protein is eluted off. The elution can also be treated with Proteinase K to remove the protein, leaving the RNA it was bound to. To analyze what species are present in the RNA elution, the sample can be reverse transcribed

and assayed via qPCR. The protein sample can be analyzed via Western blot. Treated and untreated samples can be compared to assess the relative amount of protein or RNA of interest.[28] Figure 2 shows a schematic of an IP.

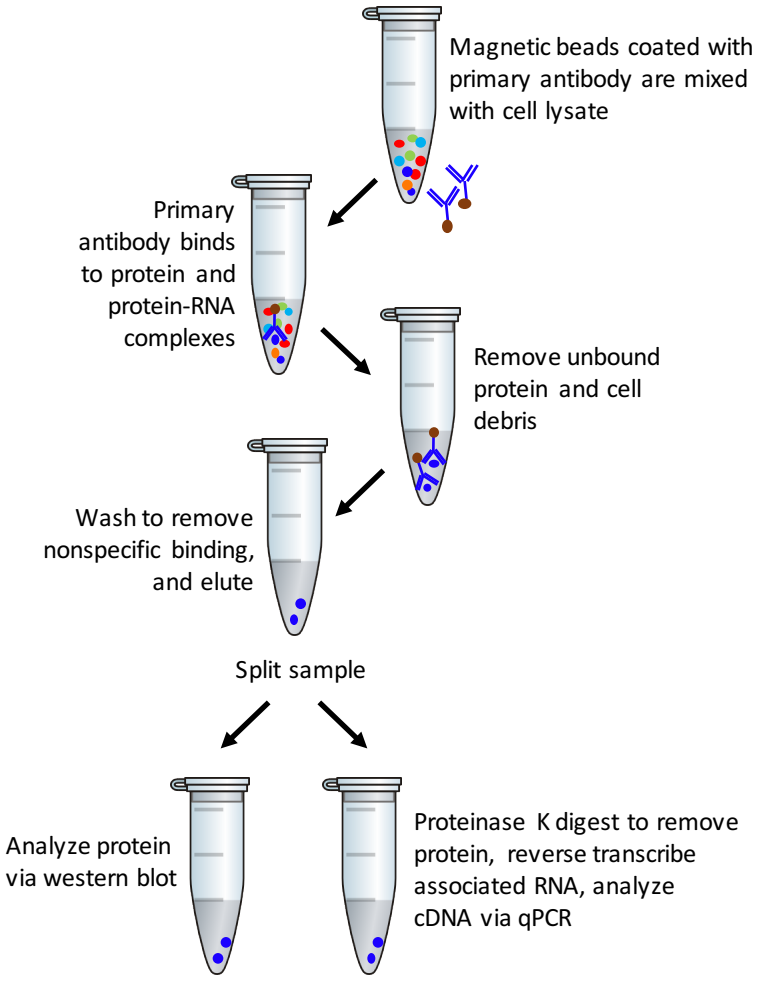


Figure 2.2 General schematic showing the steps of an immunoprecipitation experiment.

Figure 3 shows the results of one IP I did, detecting the captured protein via western blot. Figure 4 shows the fold change expression of GRAS1 measured via qPCR from the same IP experiment.

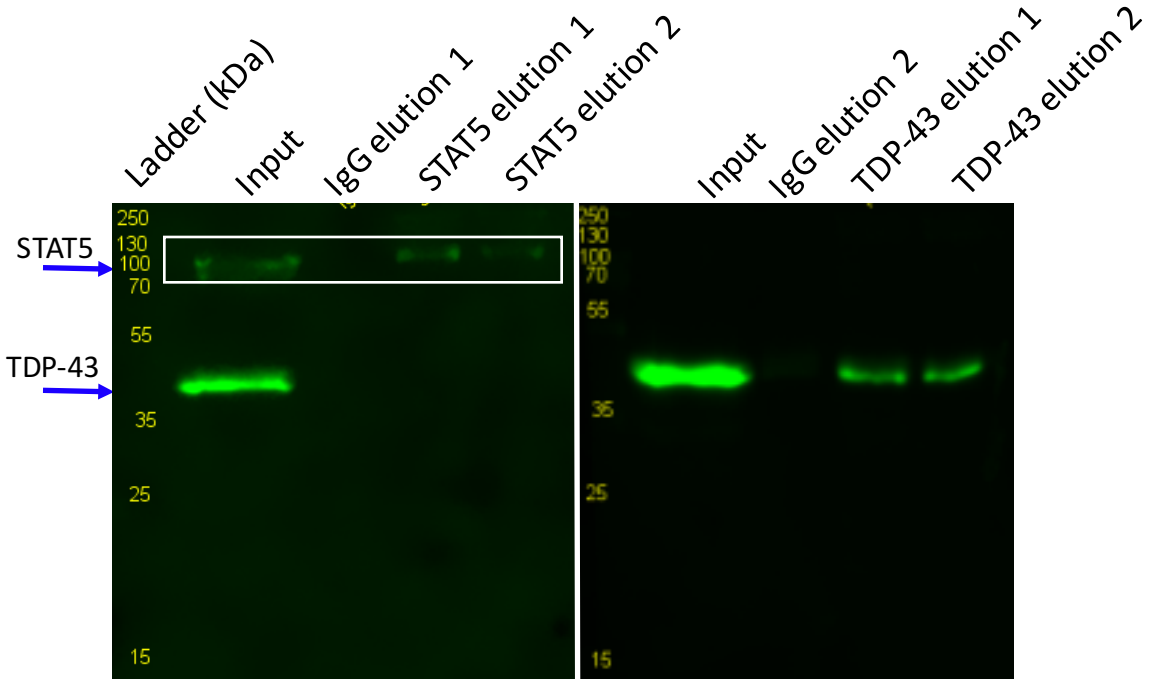


Figure 2.3 Western blot showing the detection of STAT5 and TDP-43 following an immunoprecipitation.

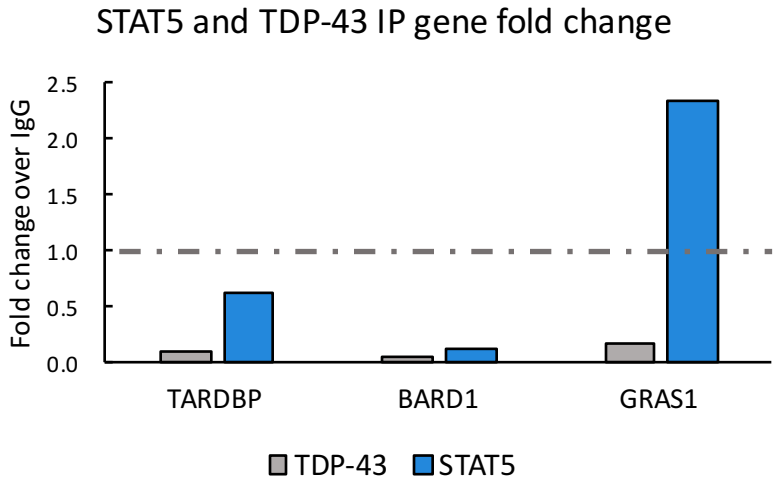


Figure 2.4 Bar graph showing the fold change expression of TARDBP, BARD1, and GRAS1 following an immunoprecipitation of STAT5, TDP-43, and IgG.

Following a cross-linked immunoprecipitation, the RNA associated with the captured protein can be analyzed to provide information on the binding sites. To do this, cells are first cross-

linked with UV light to stabilize the RNA-protein interactions, then lysed, and treated with a limited RNAse 1 digest to fragment the RNA.[29] The sample is fragmented to produce smaller pieces, which is important for sequencing later. The RNA-protein complexes are then extracted from the cell lysate by immunoprecipitation. Stringent washes are important to ensure that any nonspecific binding of random protein to the antibody is not included in the sample. Next, the RNA fragments are dephosphorylated, and an in-line barcode adapter is added to the 3' end. The sample is run on a protein gel using electrophoresis and transferred to a nitrocellulose membrane. An area 75 kDa, corresponding to about 220 nt of RNA, is extracted from the membrane to collect the captured protein and any bound RNA fragments it's crosslinked to.[29] The sample is then treated with Proteinase K to remove the bound protein and isolate the RNA. The RNA fragments are prepared for high-throughput sequencing libraries, and reverse transcribed to produce cDNA. Reverse transcription will terminate at the cross-linked site. Following reverse transcription, the previously added barcode on the 3' end of the RNA fragment is now on the 5' end of the cDNA, and another adapter that contains a random-mer sequence is ligated to the 3' end of the cDNA.[29] After sequencing, the reads are analyzed through a CLIP pipeline and aligned to the genome. Piles of reads from many fragments form visual "peaks", and can infer a likely binding site.

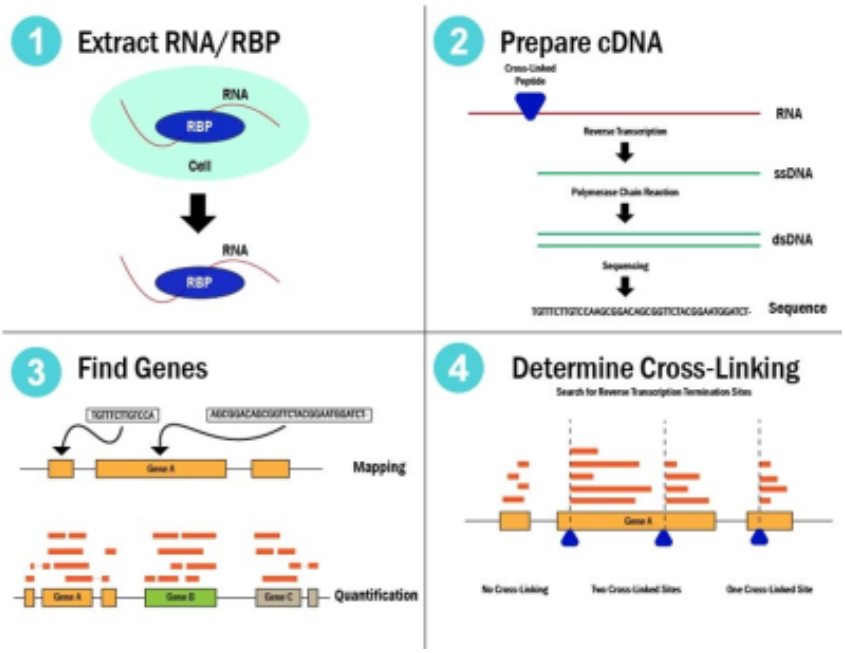


Figure 2.5 General scematic of iCLIP data analysis. Created by Jonathan Zhu.

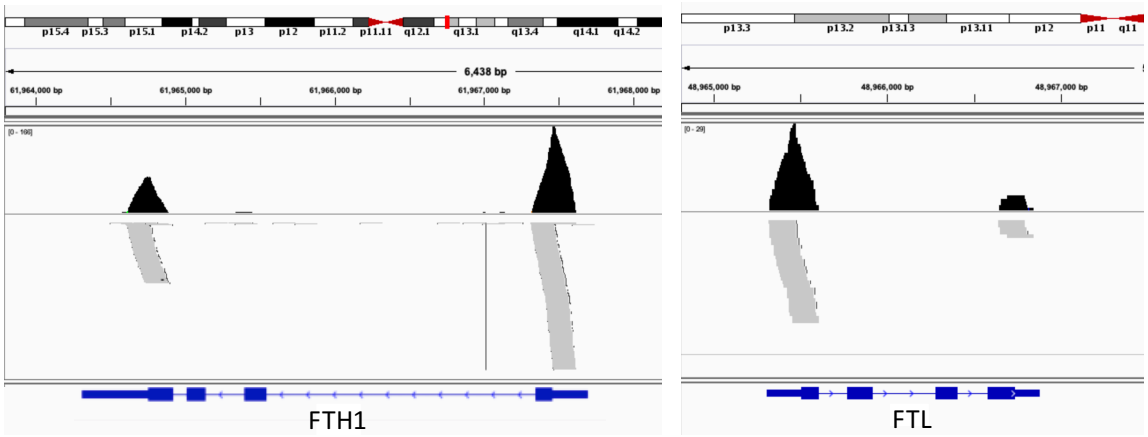


Figure 2.6 Sequencing reads from the Lin Lab iCLIP align to FTH1 and FTL.

I selected six genes that were differentially expressed by GRAS1 knockdown or overexpression and that were found to be strong binding targets of NKAP based off the CLIP analysis. I compared where each gene was bound to NKAP by visualizing which exons of the genes had peaks in Integrative Genomics Viewer (IGV), and designed qPCR primers that would target those areas. I wanted to validate if the NKAP target-genes were up or down regulated by

GRAS1 expression in A549 cells. This step was important because qPCR only amplifies a relatively short section, approximately 200 nucleotides, so if the primers didn't amplify that section of the gene, it could give a false negative of which areas were being differentially expressed by GRAS1 knockdown. I performed qPCR primer efficiency experiments on each of the six NKAP-target genes. One example of a primer efficiency test is shown in Figure 2.7.

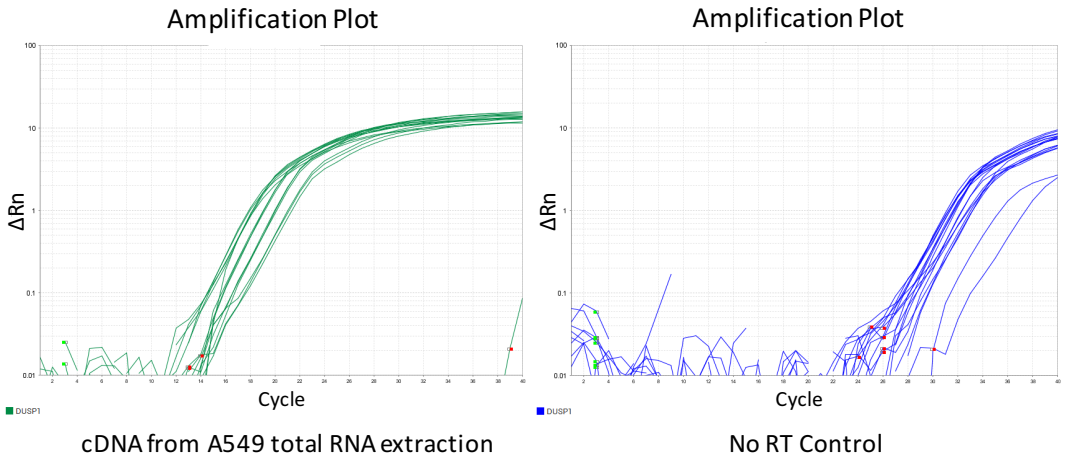


Figure 2.7 Amplification plots of NKAP targets showing effective qPCR primers.
(Left) Amplification plot showing the qPCR primer efficiency of DUSP1 qPCR primers with cDNA produced from A549 cell RNA. (Right) Amplification plot showing the qPCR primer efficiency of DUSP1 qPCR primers with a template of identical concentrations as Figure 2.7A that was not reverse transcribed.

After validating the primer efficiencies, I learned how to culture and transfect A549 cells so I could perform GRAS1 knockdown experiments using two custom GapmeRs that Tong designed. Unfortunately, the GapmeRs take a long time to produce, and we ran out of our stock before the replacement arrived. While I was waiting for the GapmeRs so I could do more biological replicates, I decided to try a GRAS1 siRNA that hadn't yet been used in A549 cells in our lab. The results of the siRNA knockdown efficiency are shown in Figure 2.8.

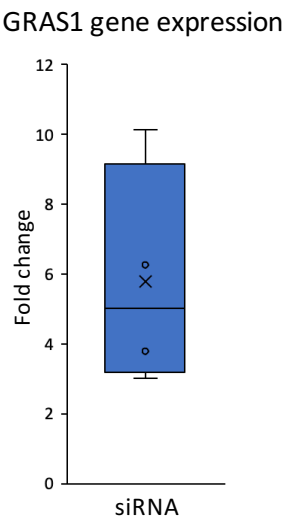


Figure 2.8 Boxplot of GRAS1 gene expression after transfecting with GRAS1 siRNA in A549 cells measured by qPCR.

Figure 2.9 shows the expression fold change of two NKAP target genes after knocking down GRAS1 in A549 cells.

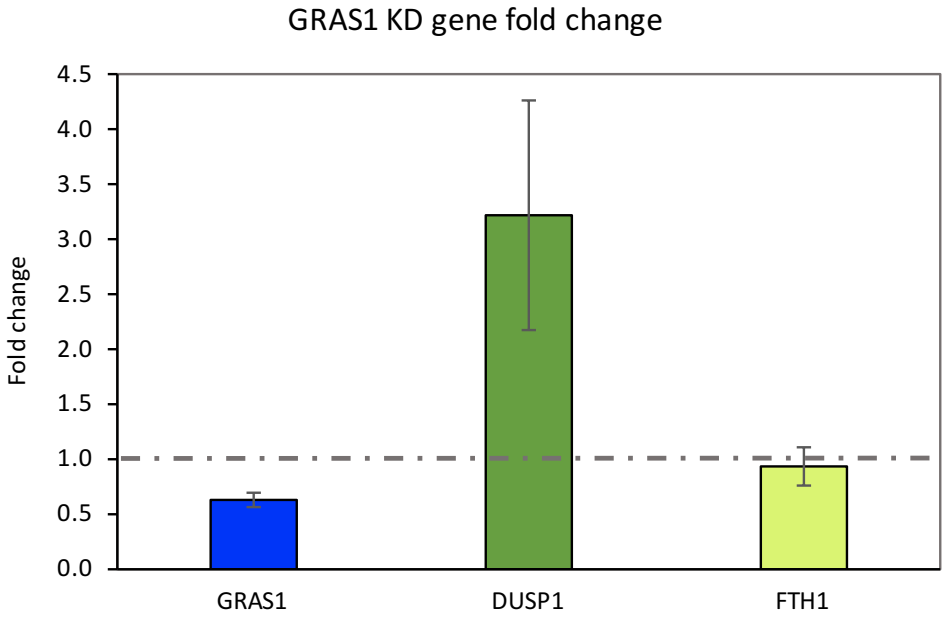


Figure 2.9 GRAS1 gene expression knockdown in A549 cells measured by qPCR.

Conclusions

Although the reason remains unclear, the GRAS1 siRNA increased expression of GRAS1 lncRNA in A549 cells. Following knockdown of GRAS1 via GapmeR, the expression of DUSP1 increased, and expression of FTH1 decreased. The other NKAP target genes were inconsistently up or down regulated. I hypothesize the inconsistencies were due to variables in the GRAS1 knockdown experiments. Because a decrease in GRAS1 leads to a loss in cell viability, sometimes too many of the A549 cells would die before the cells could be collected and lysed, thus affecting the amount of total RNA available to extract. This was particularly seen when gapmeR1 was used.

Acknowledgements

Chapter 2, in part, is a reprint of the material as it appears in bioRxiv . Su, Tong; Trang, Nu; Zhu, Jonathan; Kong, Lingbo; Cheung, Darin; Chou, Vita; Ellis, Lauren; Camden, Nichelle; Huang, Calvin; McHugh, Colleen. GRAS1 non-coding RNA protects against DNA damage and cell death by binding and stabilizing NKAP. bioRxiv, 2024. The thesis author was a contributing investigator and author of this paper.

Chapter 3 NKAP Functional Assays

Introduction

Tong Su in our lab showed via RAP-MS that GRAS1 RNA interacts with NKAP. The NKAP expression levels were very low and undetectable by western blot. I was trying to improve the overexpression of NKAP to see if we could rescue the cells from cell death following GRAS1 knockdown, and to consistently overexpress NKAP so that it could be pulled out in immunoprecipitation. We wanted to pull it out to prepare samples for eCLIP experiments.

Methods

Transfection. I transfected A549 cells using Lipofectamine 2000 following the manufacturers protocol. I seeded a 12-well plate with 200,000 A549 cells and let them grow in the incubator overnight. The next day when the wells were approximately 80% confluent, I transfected the cells in each well with 1 ug of plasmid DNA mixed with 125 uL Opti-mem, and combined that mixture with 3.5 uL of Lipofectamine 2000. After incubating the DNA and Lipofectamine mixtures at room temperature for 15 minutes, I added them to respective wells by dripping small amounts using a p200 micropipette, distributing the drops evenly over the surface of the well. I grew the cells in the incubator for 48 hours. To collect the cells, I detached them using Trypsin and transferred to 1.5 mL epi-tubes. I pelleted each sample by centrifugation. I washed each pellet resuspending in 1 mL PBS, and recentrifuged. Finally, I resuspened each pellet in 140 uL of SDS Page buffer containing B-mercaptoethanol.

Western Blot. Following transfection, I harvested the cells using trypLE (Thermo Fisher Scientific) and washed them with 1x PBS. The cells were then mixed with cell lysis buffer, a protein inhibitor cocktail (539134, Millipore Sigma), and sodium fluoride. I incubated the mixture at 95°C for 10 minutes before loading 20 uL of sample per lane into a 10% page gel, and separating the proteins by SDS-PAGE. The rest of the Western Blot was done following the protocol in Chapter

2 with the following exceptions: I used electrophoresis to separate the proteins at 120V for 2 hours, and I used an NKAP antibody (ab229096, Abcam) instead of STAT5, as well as a GFP antibody (50430-2-AP, Proteintech). The same control antibodies were used as listed in Chapter 2.

Results

My first goal was to overexpress NKAP in A549 cells. I obtained a linear plasmid vector from Adarsh Balaji and planned to insert the NKAP gene using Gibson cloning techniques. Figure 1 shows the plasmid diagram and scheme for Gibson cloning.

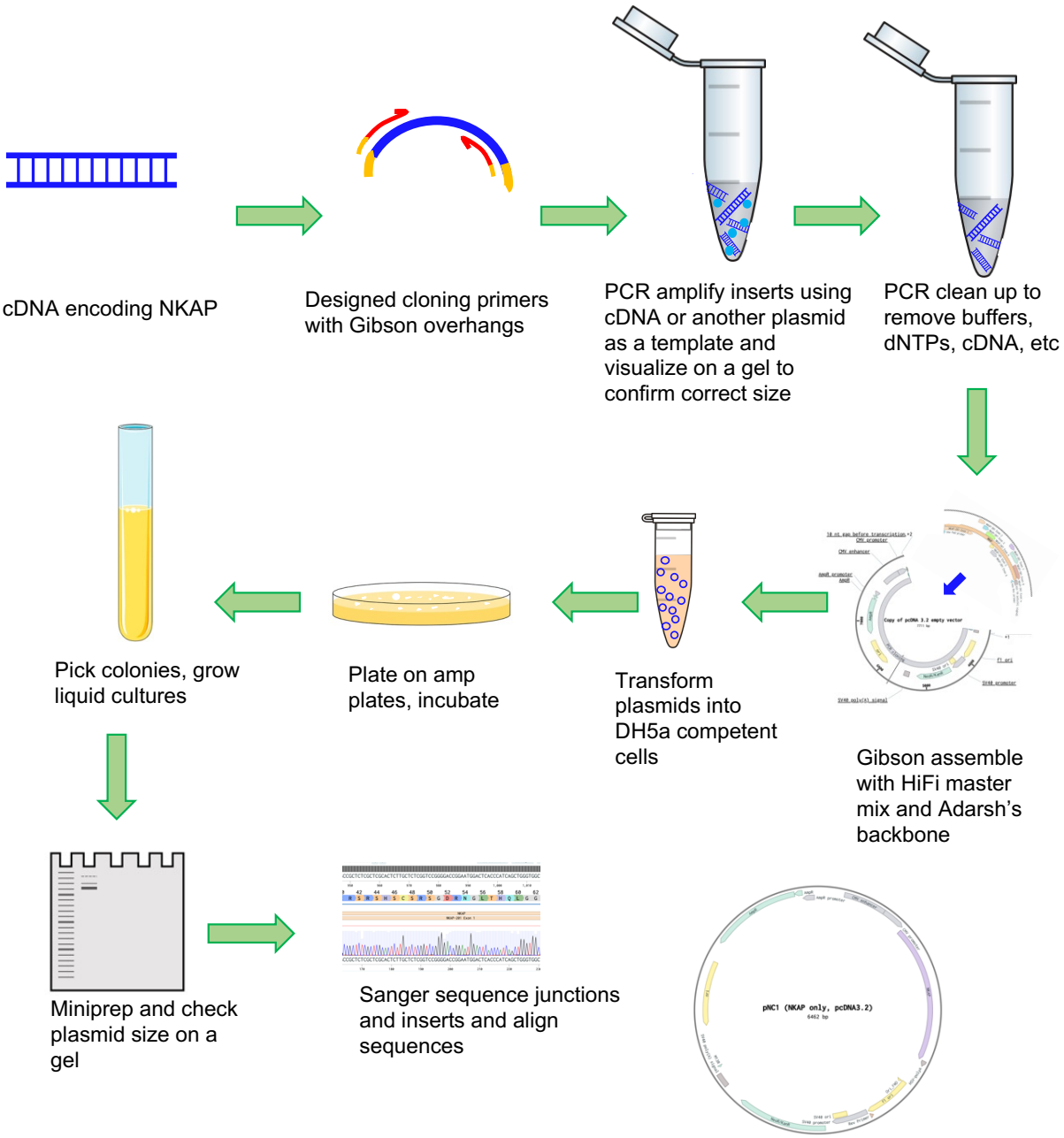


Figure 3.1 Schematic diagram of plasmid construction for NKAP expression.

Figure 2 shows the robust amplification of the NKAP gene from cDNA I produced by reverse transcribing RNA extracted from A549 cells.

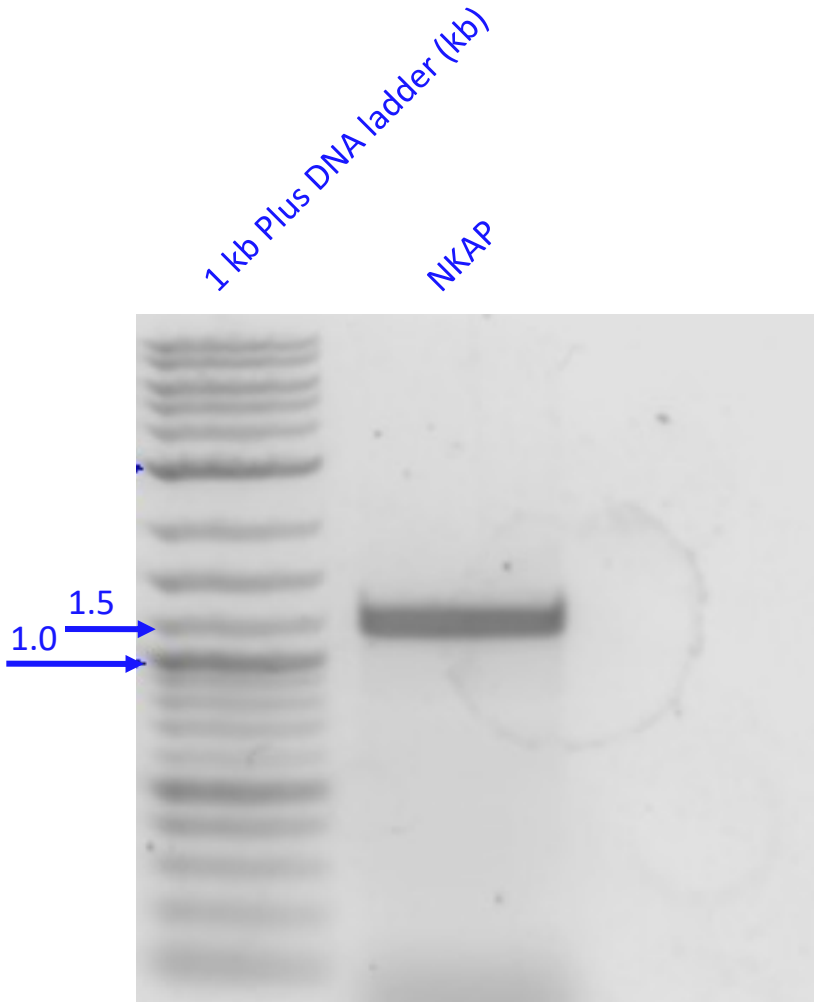


Figure 3.2 Results of PCR amplification of NKAP from cDNA

The final CMV-NKAP expression plasmids are shown in Figure 3. The plasmids were screened by restriction enzyme digest shown in Figure 4. The plasmid sequences were verified by DNA sequencing.

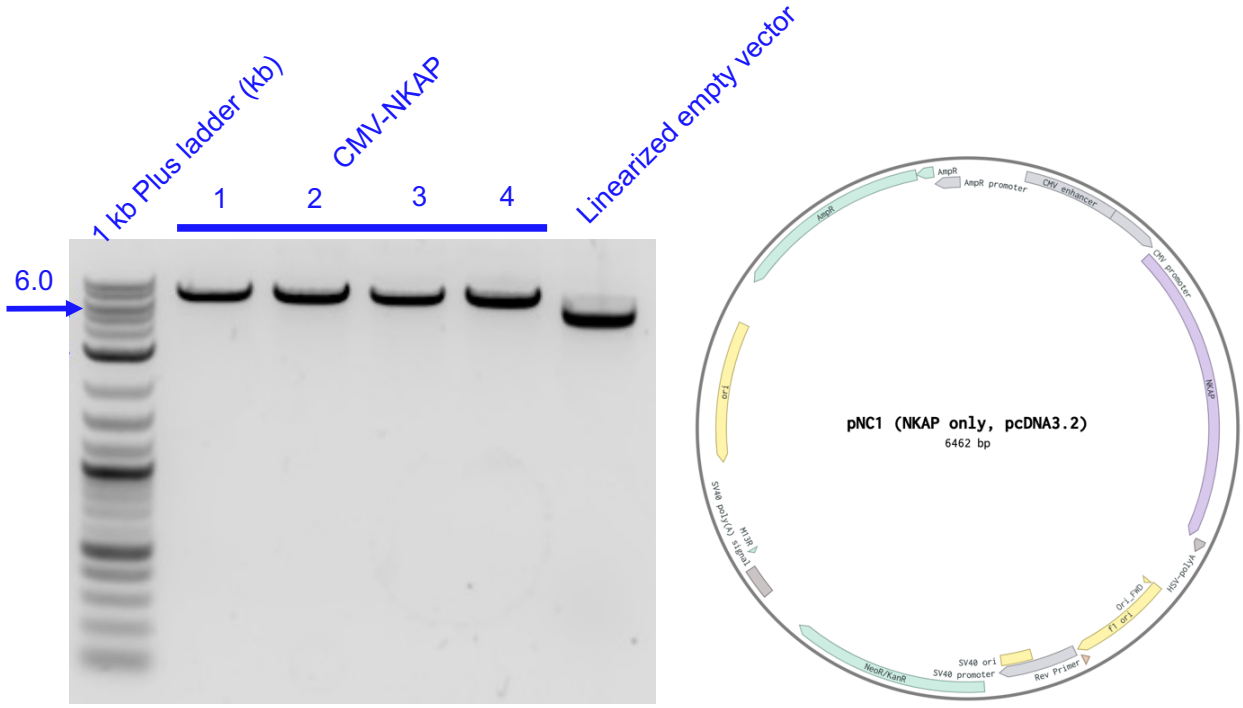


Figure 3.3 DNA gel and plasmid map of overexpression construct CMV-NKAP.

(A) DNA gel showing plasmids are the correct size and supercoiled. (B) Plasmid map of CMV-NKAP.

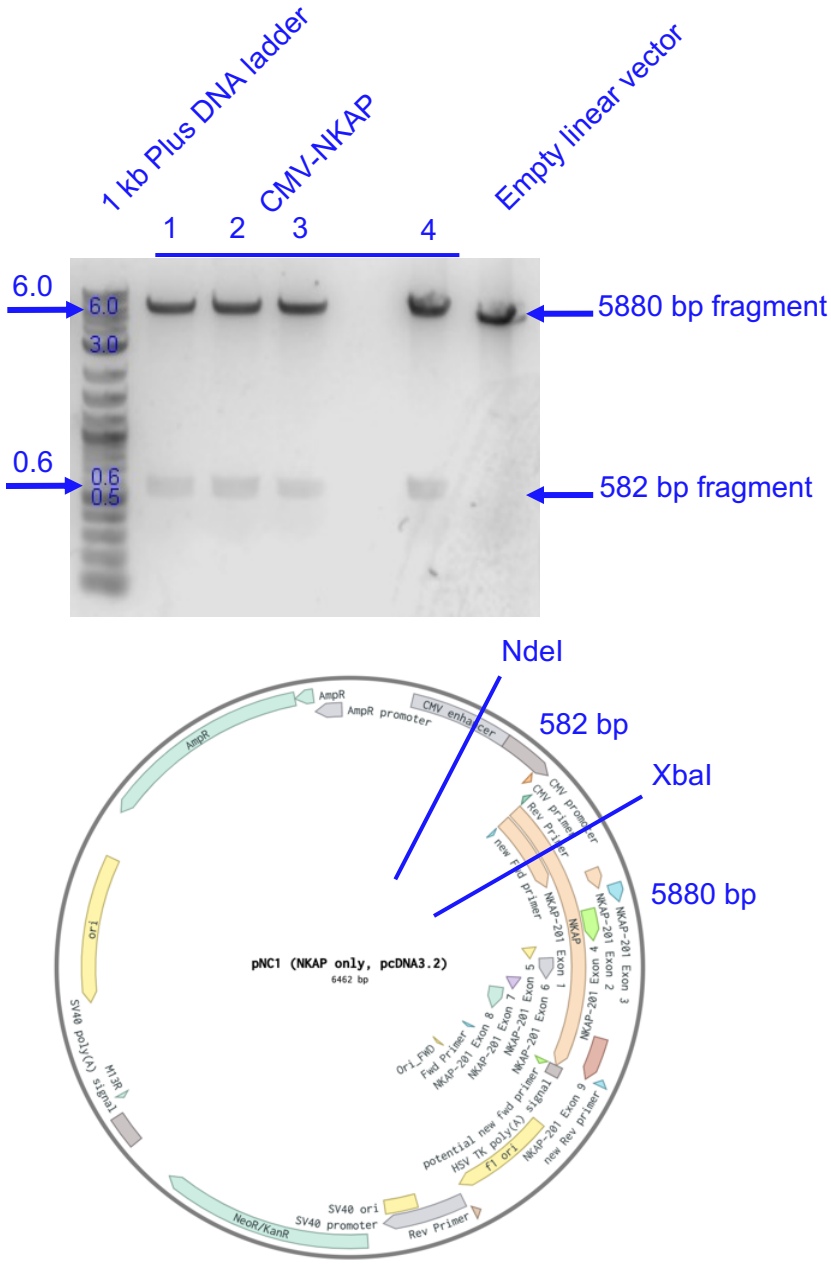


Figure 3.4 Restriction enzyme digest screening results showing successful insert of NKAP fragment.

(Top) DNA gel showing the screening results of a restriction enzyme digest using NdeI and XbaI to screen for successful insert of NKAP fragment. The lane between 3 and 4 was not used due to junk inside. (Bottom) Plasmid map of CMV-NKAP showing the cut sites of NdeI and XbaI.

I transfected A549 cells and attempted to detect NKAP using different antibodies, but detection was very poor. I decided to generate a new expression plasmid that would produce Myc-tagged NKAP for ease of detection. To do this, my plan was similar to the one diagramed in

Figure 3.1 using Gibson cloning to insert both the Myc tag and NKAP gene. Figure 3.5 shows the PCR optimization for production of the inserts with Gibson overhangs.

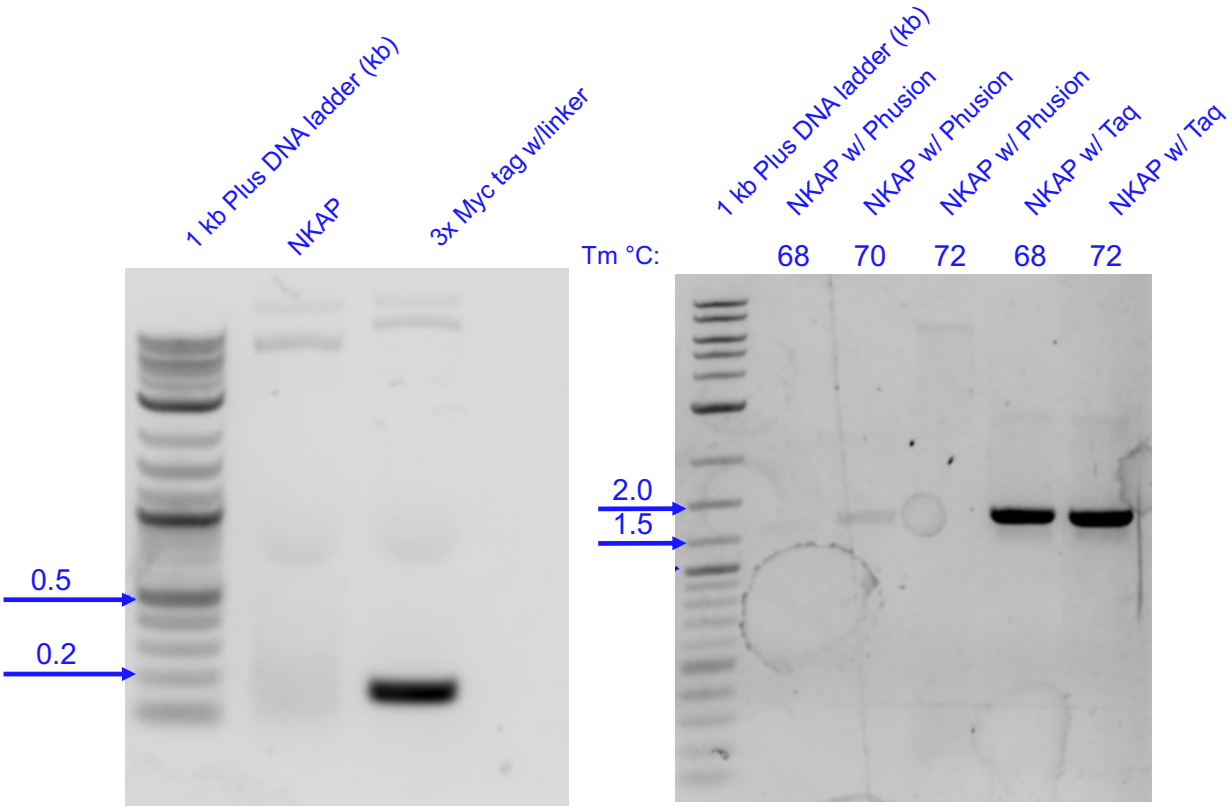


Figure 3.5 PCR amplification of 3x Myc tag and optimization of NKAP PCR amplification. (Left) PCR amplification of 3x Myc tag for construction of Myc-NKAP expression plasmid. (Right) Optimization of NKAP PCR amplification.

The final CMV-Myc-NKAP expression plasmids are shown in Figure 3.6. The sequences were verified by DNA sequencing.

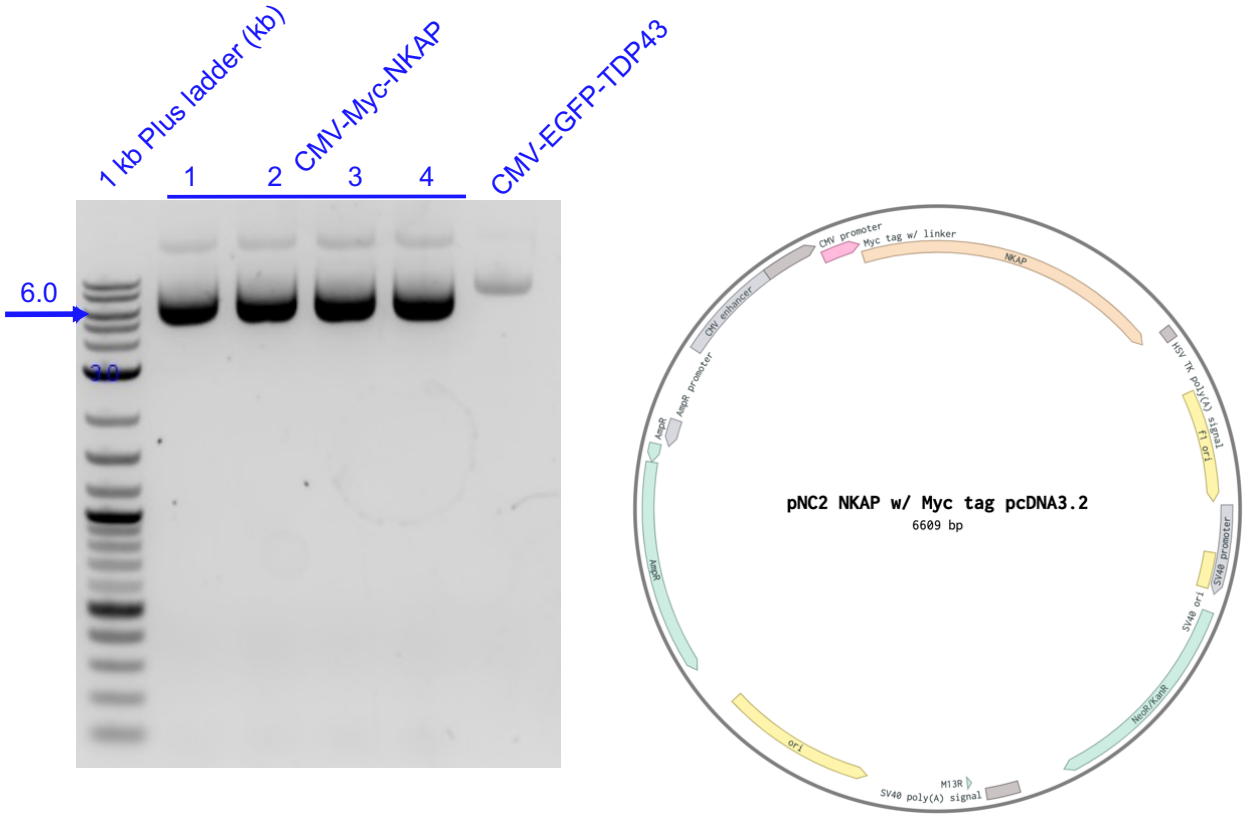


Figure 3.6 DNA gel and plasmid map of overexpression construct CMV-Myc-NKAP.
(A) DNA showing the CMV-Myc-NKAP expression plasmids are the correct size. (B) Plasmid map of CMV-Myc-NKAP.

Unfortunately, after transfecting A549 cells with Myc-tagged NKAP expressing plasmid, I was unable to detect the protein on western blot using the available Myc antibodies. My next plan was to create an EGFP-fusion construct that would create a fluorescent NKAP, avoiding the need for western blots. Using a similar scheme to that shown in Figure 3.1, I created two new plasmids: a fusion construct, and an EGFP control that did not contain NKAP.

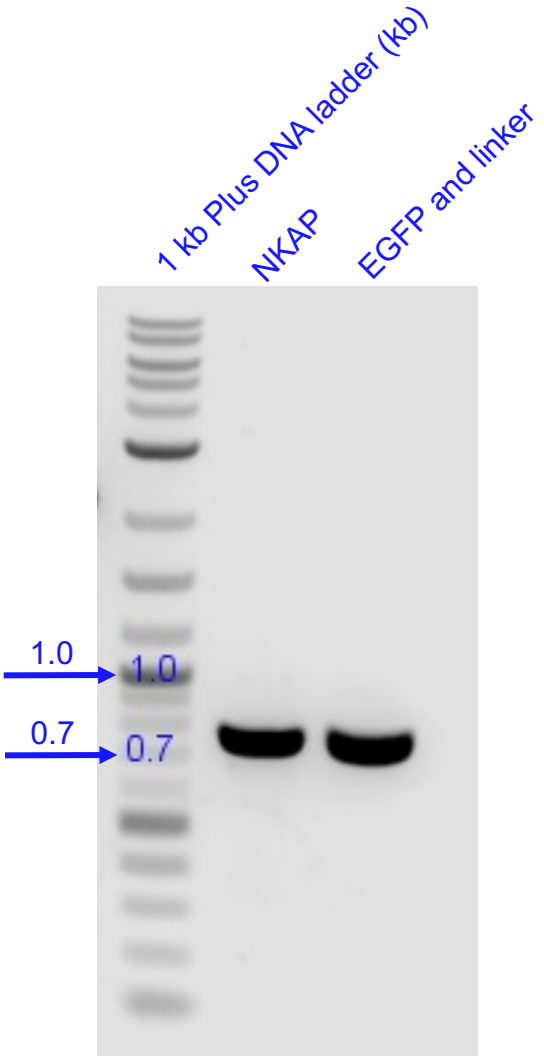


Figure 3.7 DNA gel showing the PCR amplification of NKAP and EGFP for construction of NKAP-EGFP fusion and EGFP control expression plasmid.

Figure 3.7 shows PCR amplification of the NKAP and EGFP inserts for the fusion and EGFP control expression plasmids. The final CMV-EGFP-NKAP and CMV-EGFP expression plasmids are shown in Figure 3.8. The sequences of both plasmids were verified by DNA sequencing.

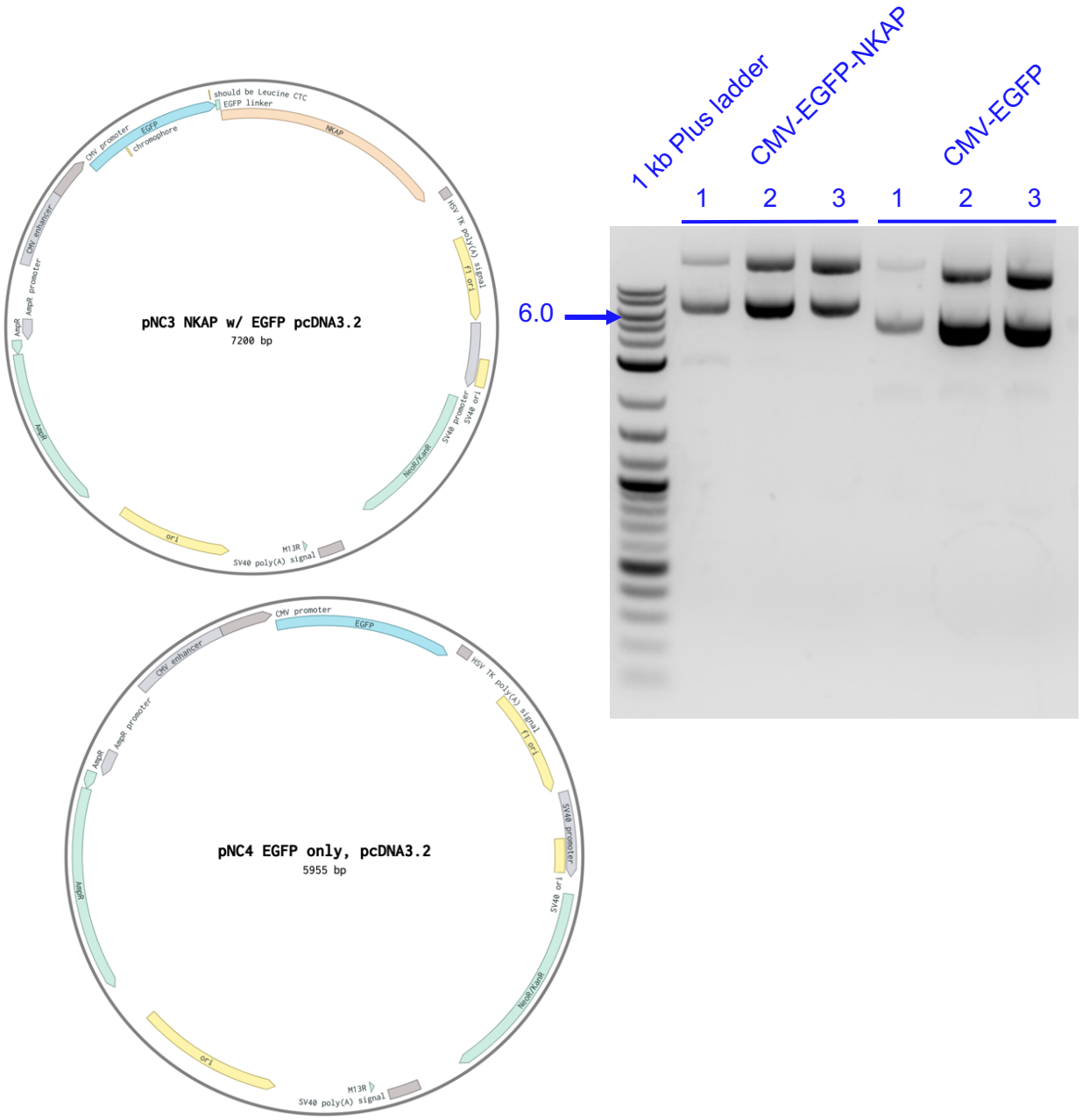


Figure 3.8 Plasmid map and verification of NKAP overexpression contruct and EGFP control. (Right) DNA gel showing plasmids are the correct size and supercoiled. (Top left) Plasmid map of CMV-EGFP-NKAP. (Bottom left) Plasmid map of CMV-EGFP.

Following a transfection in A549 cells with the new plasmids, I collected each sample and ran them on a western blot to probe for NKAP and EGFP. Unfortunately, it was difficult to discern true green cells from noise using a Bright-field microscope due to autofluorescence and different

microscope settings. Figure 3.9 shows Brightfield microscope images of A549 cells transfected with CMV-EGFP, CMV-EGFP-NKAP, or CMV-Myc-NKAP. It is common for cells to autofluoresce following a transfection, and by adjusting the settings on the microscope, I mistook sick, green autofluorescent cells for a positive confirmation that my transfection was successful and that the cells were producing the new proteins. However, confirmation via western blot was very low. I tried many different amounts and combinations of DNA and lipofectamine2000 to optimize the transfection. Table 1 shows an example of one optimization experiment. After many attempts, I could not detect NKAP using the available antibody, and could only detect a faint band for EGFP, but only when I concentrated the samples. Figure 3.10 shows an example of a western blot following one optimization experiment, corresponding to Table 1.

Table 1 Optimizing transfecting EGFP-containing constructs into A549 cells in a 6-well plate.

Plasmid	CMV-EGFP-NKAP	CMV-EGFP-NKAP	CMV-EGFP-NKAP	CMV-EGFP-NKAP	CMV-NKAP	CMV-EGFP
DNA amount (ug)	2.5	1.2	2.5	1.2	2.5	2.5
Lipofectamine2000 (uL)	5.0	5.0	7.5	7.5	7.5	7.5

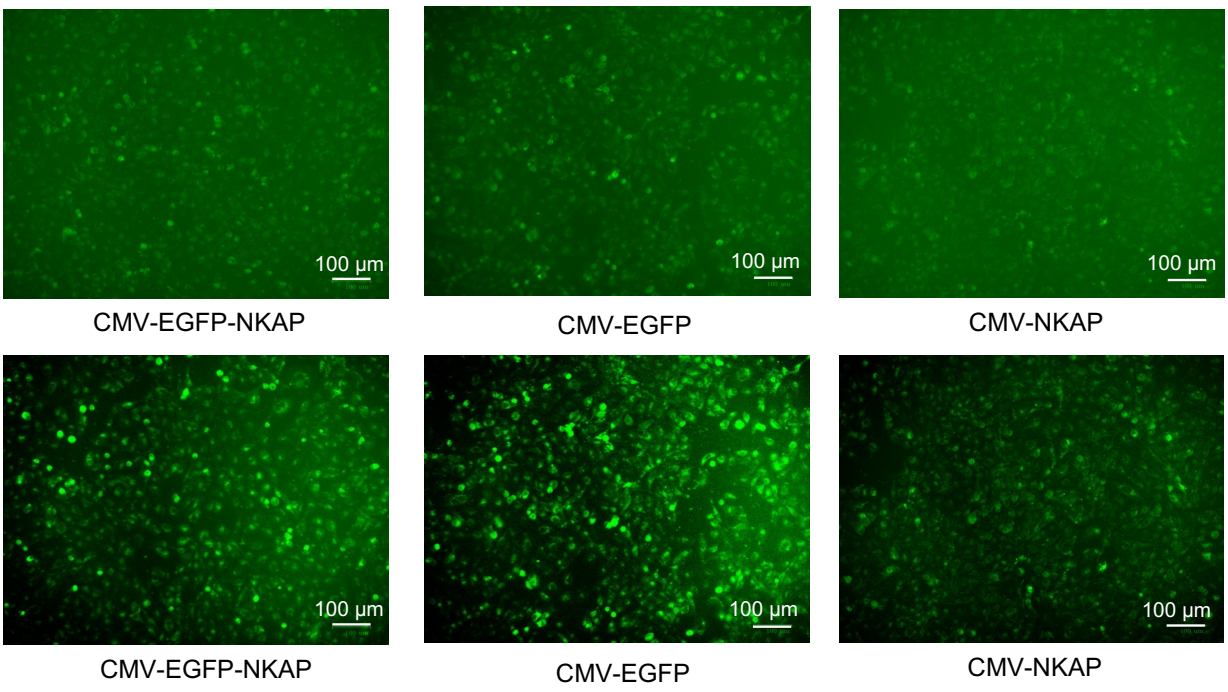


Figure 3.9 Microscope images of A549 cells taken in the blue channel after transfecting with CMV-EGFP, CMV-EGFP-NKAP, or CMV-NKAP. Adjusting the exposure, brightness, and contrast setting in an attempt to visualize what I thought were faintly expressing EGFP cells produced false positives.

While the cause of autofluorescence is unclear, one likely hypothesis is that it's due to the presence of NADH (nicotinamide adenine dinucleotide (reduced to include) Hydrogen), a coenzyme involved in energy production, which has broad excitation and emission spectra similar to the blue light channel used to detect EGFP. EGFP absorbs light at 488 nm and emits light at 507 nm, NADH absorbs light around 340-360 nm and emits around 430-450 nm.[30] Although outside of its peak excitation spectrum, the blue light can cause NADH to fluoresce faintly enough to be confused with faintly expressing EGFP. Additionally, adjusting the exposure, brightness, and contrast settings on the microscope can contribute to mistaking autofluorescence as a false positive. Because I cloned the EGFP and linker fragment from another plasmid used in the McHugh lab that was confirmed to work, and verified that its sequence is correct, I initially thought my cells were successfully expressing EGFP.

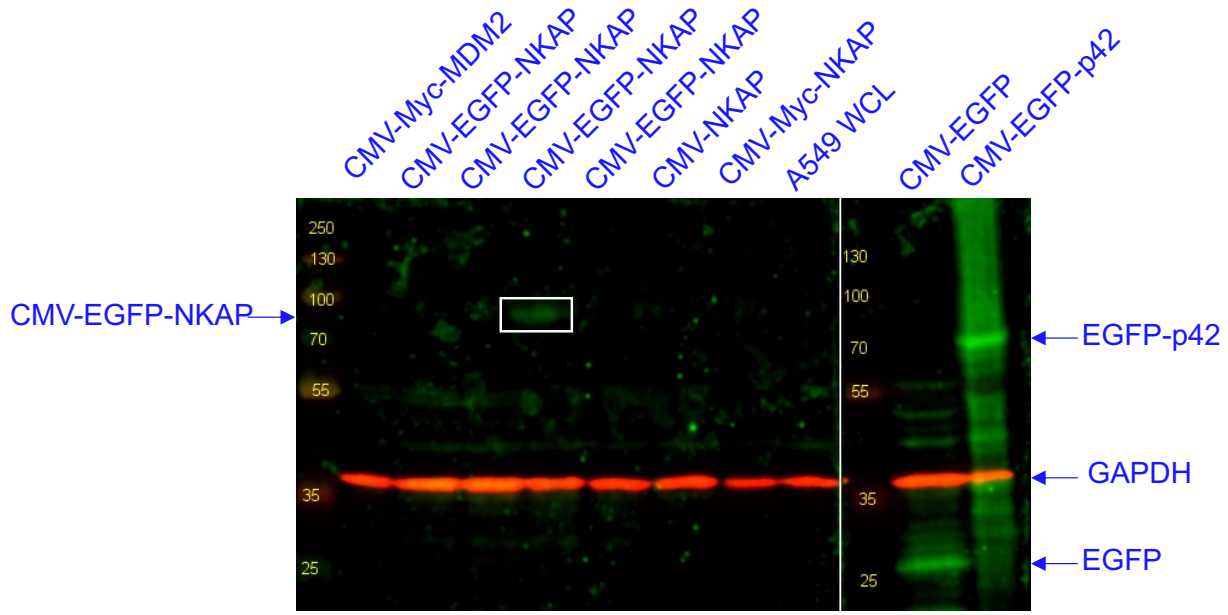


Figure 3.10 Western blot of A549 cell lysate following a transfection, probed for EGFP.

Lanes 3-7 and 11 are samples from an optimization experiment and contain approximately 200,00 A549 cells, approximately 2.5x as many cells as I would typically load in a sample. Lanes 2, 8, 9, and 12 are controls.

To determine if there was an increase of NKAP DNA being expressed, but not translated, I used qPCR to calculate the fold change of NKAP gene expression in my samples, shown in Figure 3.11 and Table 2. My qPCR data indicated an increase in NKAP expression, but the lack of green cells visible through the microscope, and lack of detection via western blot, suggested the protein was not being translated correctly.

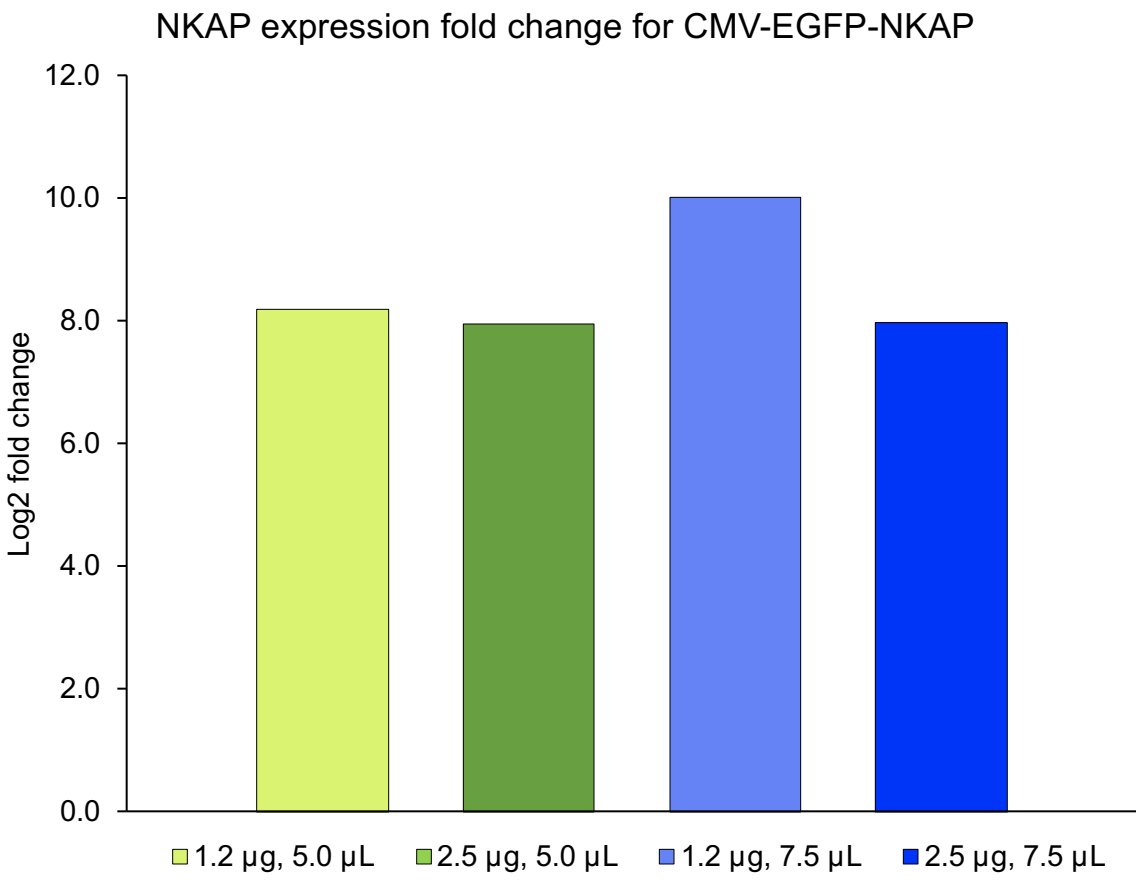


Figure 3.11 Transfection of NKAP overexpression plasmids increases NKAP mRNA.

Bar graph showing the increase of NKAP gene expression following transfecting A549 cells with 1.2 µg or 2.5 µg of CMV-EGFP-NKAP overexpression plasmid and 5.0 µL or 7.5 µL of lipofectamine2000, measured by qPCR.

Table 2 NKAP gene fold change expression values of CMV-EGFP-NKAP measured by qPCR.

Each sample was transfected with either 1.2 µg or 2.5 µg of overexpression vector DNA and 5.0 µL or 7.5 µL of Lipofectamine2000.

Gene expression fold change			
	DNA (µg)	Lipofectamine2000 (µL)	NKAP
CMV-EGFP-NKAP	1.2	5.0	8.183
	2.5	5.0	7.954
	1.2	7.5	10.01
	2.5	7.5	7.978

Due to evidence that the proteins in my over expression vectors were not being translated correctly, I decided to reexamine the design of my plasmid constructs. I found that my expression vectors did not contain a Kozak consensus sequence, which has been shown to be important for

initiating ribosome recognition and assembly, and therefore translation.[31] I also realized that my plasmids were unnecessarily large due to containing extra components, such as kanamycin resistance, that wouldn't be useful in my experiments, but could contribute to lower transfection efficiency due to increasing the overall size of the plasmid.[32] For these reasons, as well as time constraints, I decided to use a different vector backbone instead of cloning in a Kozak sequence and using the same linear backbone that I had used for my previous plasmids.

I obtained three overexpression plasmids that were known to express well via transient expression in multiple cell types from the Komor lab[33]. I used them in a transfection experiment with A549 cells to visualize how well they expressed EGFP in the cell type, and to ensure that I could successfully transfect cells.

Two of the plasmids contained the CMV promoter and identical backbones, and another contained the pMAX promoter. I was told that the pMAX plasmid would have very high expression, so I included it as a control for transfection. I wanted to have a backup in case the others didn't express as well, even though I didn't plan to use the pMAX promoter in my overexpression vectors. I also included another EGFP-containing plasmid, pcEGFP (CMV-GFP), as a control that others in my lab have used. I hadn't had success previously with pcEGFP, but I wanted to try it again and compare the expression with the others. Figure 3.12 shows microscope images of the results of this transfection

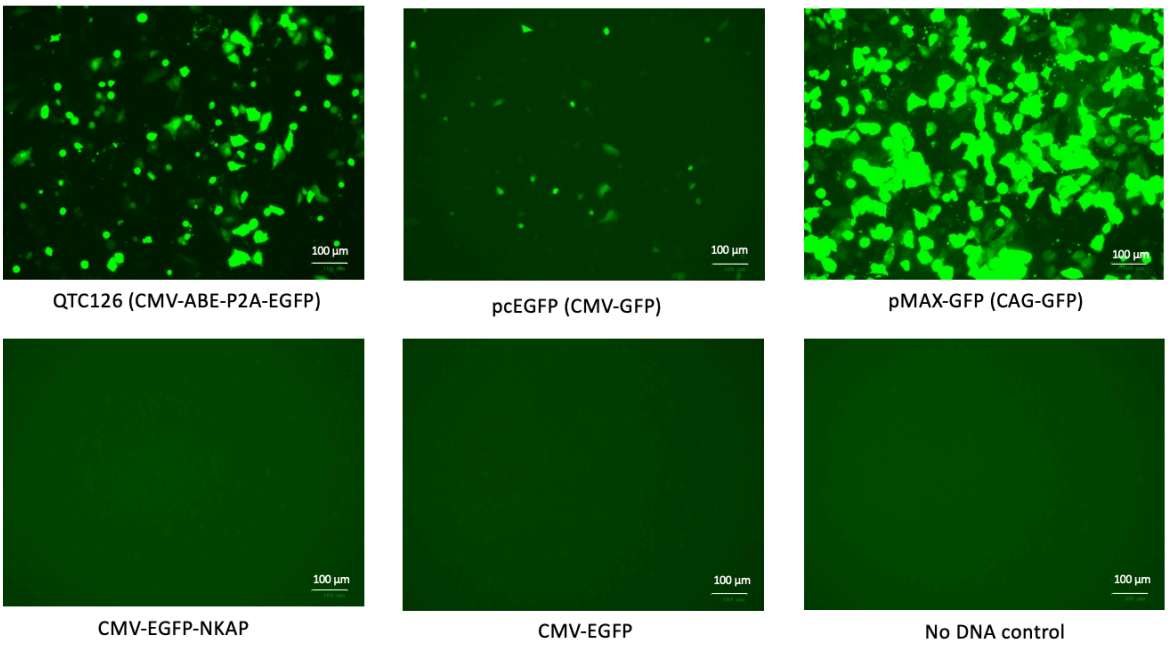
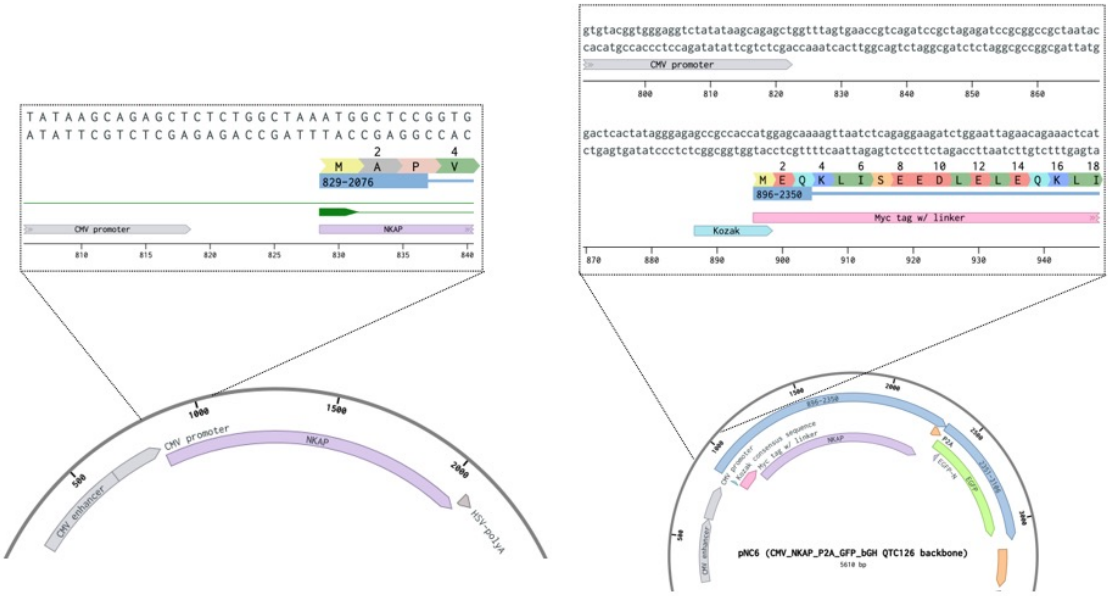


Figure 3.12 Microscope images of A549 cells transfected with the expression vectors CMV-ABE-P2A-EGFP, pcEGFP, pMAX-GFP, CMV-EGFP-NKAP, or CMV-EGFP and Lipofectamine2000 following the manufacturer's protocol. All the images were taken with the microscope's recommended standard settings for the blue light channel, with the exception of the pMAX sample in which the brightness had to be decreased in order to visualize cells clearly. No EGFP expression was detected for my CMV-EGFP-NKAP and CMV-EGFP constructs.

Next, I designed new plasmids to overexpress NKAP. I wanted one of them to incorporate a P2A-EGFP component, so that I could express EGFP in the same cells without having to co-transfect a separate plasmid, but still have a positive control. P2A is a viral polycistronic sequence that causes ribosomal skipping, which allows two peptides to be translated in the same open reading frame without creating a fusion construct.[34] Fortunately, one of the borrow plasmids already contained such a sequence, so I was able to amplify it out using PCR, and ligate it to an NKAP fragment. I also wanted to create another EGFP-NKAP fusion construct. To clone these plasmids, I used the same method outlined in Figure 3.1. Figure 3.13 highlights the addition of the Kozak sequence in the new plasmid designs.



pNC1: CMV_NKAP_HSV-polyA

pNC6: CMV+Kozak_3xMyc_NKAP_ P2A-GFP_bGHR-polyA

Figure 3.13 Two plasmid maps highlighting the addition of the Kozak sequence in the design for CMV-Myc-NKAP-P2A-EGFP and CMV-EGFP-NKAP. The latter is not shown in this figure, but does contain the Kozak sequence.

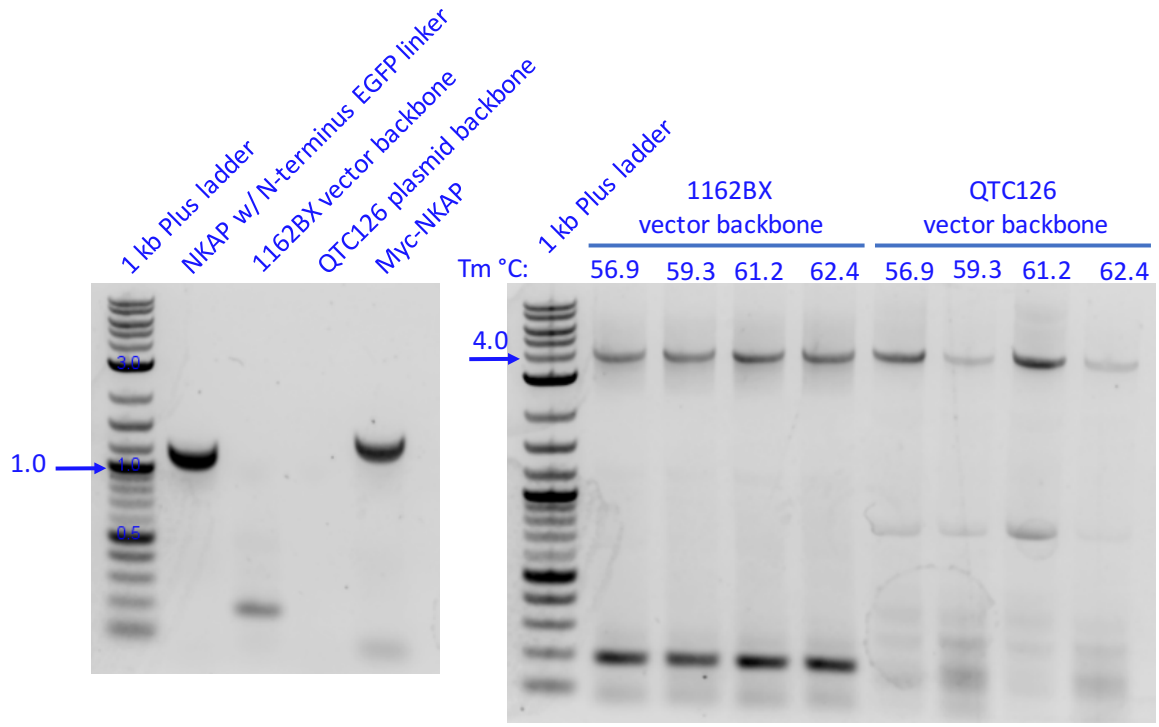


Figure 3.14 Successful PCR amplification of NKAP, Myc-NKAP, and P2A plasmid backbones. (A) PCR amplification of NKAP with an N-terminus EGFP linker and a 3x Myc tag for construction of EGFP-NKAP and Myc-NKAP-P2A-EGFP expression plasmids, respectively. (B) Optimization and PCR amplification of the vector backbone to the plasmids 1162BX (CMV-EGFP-bGH) and QTC126 (CMV-ABE-P2A-EGFP-bGH).

Figure 3.14 shows PCR amplification of the NKAP and 3x Myc inserts, as well as both vector backbones for the EGFP-NKAP and Myc-NKAP-P2A-EGFP expression plasmids. Figure 3.15 shows a DNA gel screening for the correct size of the plasmids prior to sequencing. The final EGFP-NKAP and Myc-NKAP-P2A-EGFP expression plasmids are shown in Figure 3.16. The sequences of both plasmids were verified by DNA sequencing.

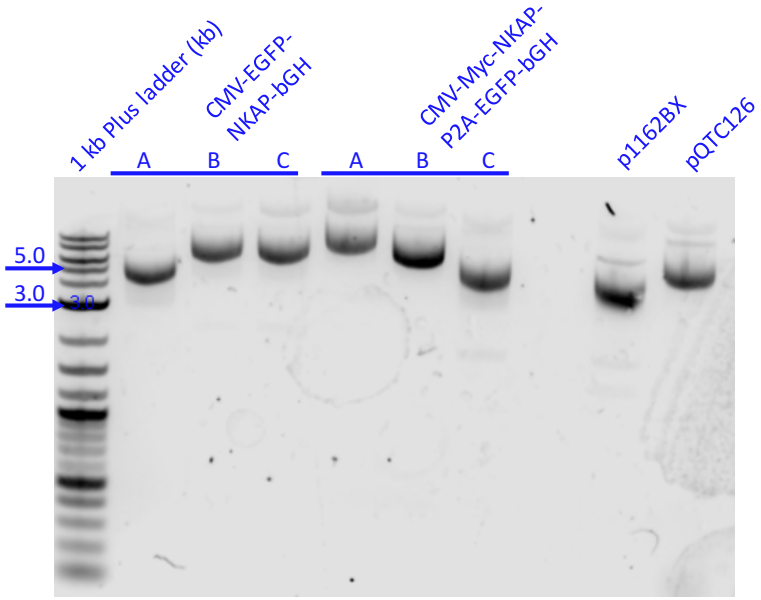


Figure 3.15 DNA gel screening plasmids for the correct size before sequencing.

The plasmids 1162BX (CMV-EGFP-bGH) and QTC126 (CMV-ABE-P2A-EGFP-bGH) were ran as a size comparison. The B replicates for both the CMV-EGFP-NKAP-bGH and CMV-Myc-NKAP-P2A-EGFP-bGH plasmids were selected for sequencing and verified to be correct.

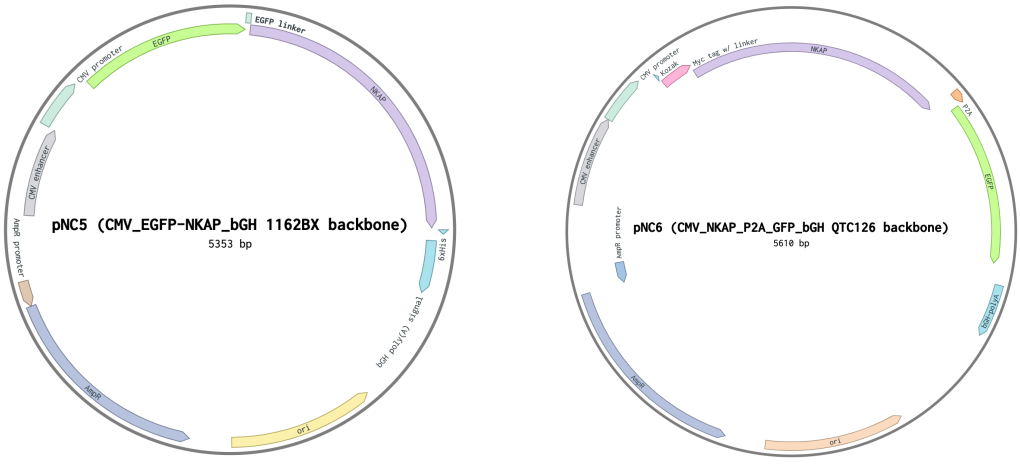
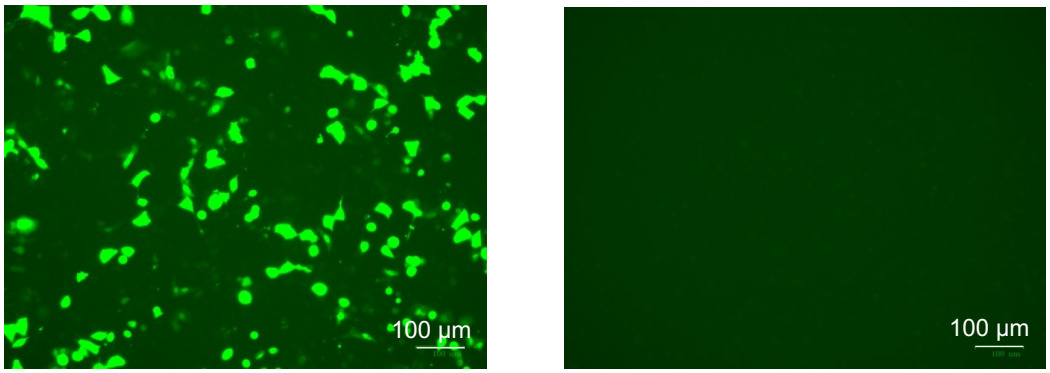


Figure 3.16 Plasmid maps of CMV-EGFP-NKAP-bGH and CMV-Myc-NKAP-P2A-EGFP-bGH (Left) Plasmid map of CMV-EGFP-NKAP-bGH. (Right) Plasmid map of CMV-Myc-NKAP-P2A-EGFP-bGH.

I transfected A549 cells with CMV-NKAP-P2A-EGFP or CMV-Myc-HDM2 overexpression plasmids. Figure 3.17 shows strong EGFP expression only in the A549 cells transfected with

CMV-Myc-NKAP-P2A-EGFP.



CMV-Myc-NKAP-P2A-EGFP-bGH CMV-Myc-HDM2

Figure 3.17 Comparison of EGFP expression in plasmids.

(Left) Microscope image taken in the blue channel of A549 cells transfected with CMV-Myc-NKAP-P2A-EGFP-bGH overexpression vector. (Right) Microscope image taken in the blue channel of A549 cells transfected with CMV-Myc-HDM2. Without over-exposing the field of view in the microscope, only the construct that contains EGFP fluoresces.

I attempted to detect NKAP, Myc, and EGFP via western blot using the respective antibodies. Detection using the available NKAP antibody was very poor, however, I was able to strongly detect Myc in CMV-Myc-NKAP-P2A-EGFP, and faintly in CMV-Myc-HDM2, shown in Figure 3.18. The band for CMV-Myc-HDM2 is visible when the brightness on the blot is increased.

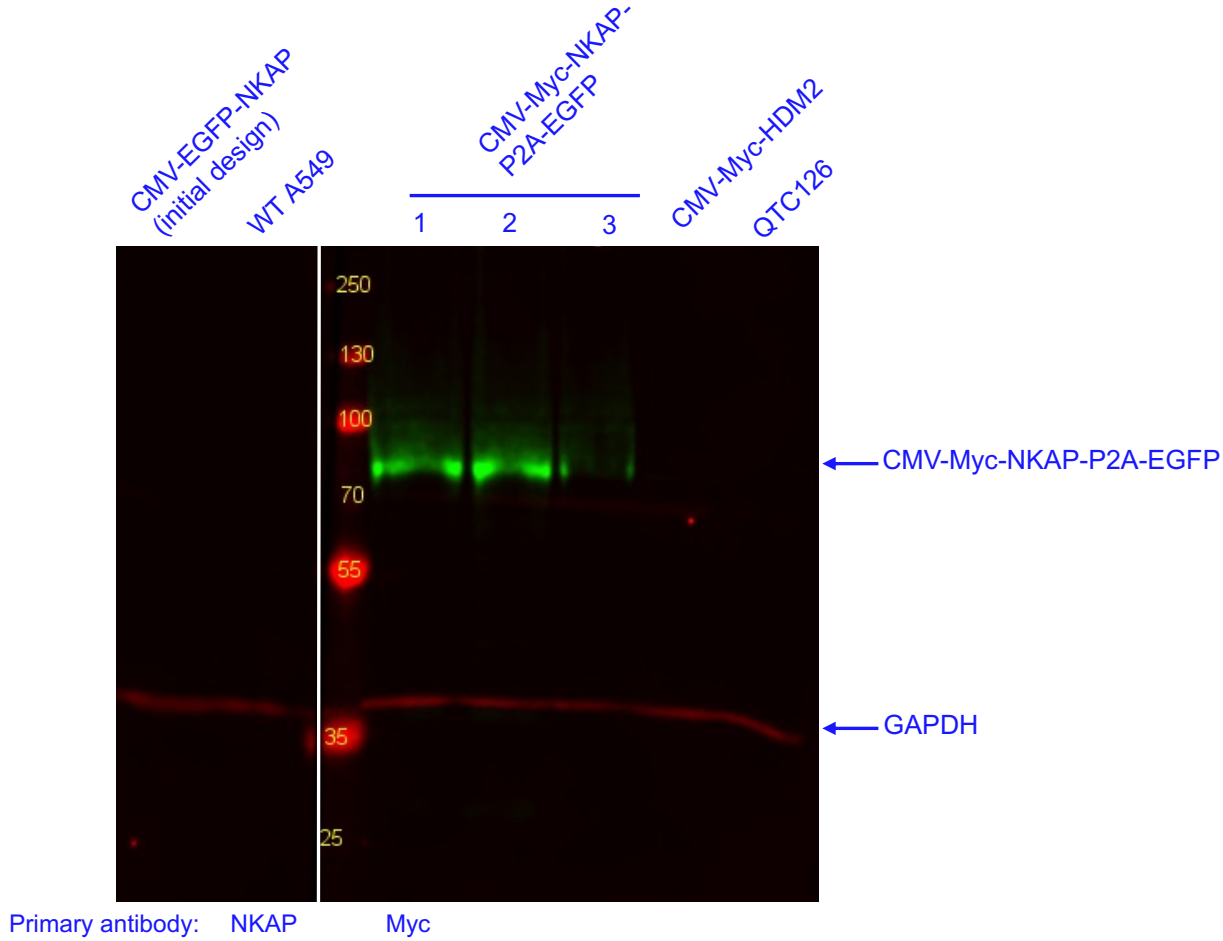


Figure 3.18 Western blot of A549 cell lysate following a transfection of CMV-Myc-NKAP-P2A-EGFP, CMV-EGFP-NKAP, CMV-Myc-HDM2, or CMV-ABE-P2A-EGFP. The blot was cut and probed for NKAP on the left and Myc on the right.

I photobleached the blot and confirmed that the bands were no longer visible between probing for different proteins. Detection of EGFP was strong and at the correct size in my CMV-Myc-NKAP-P2A-EGFP and CMV-ABE-P2A-EGFP samples, shown in Figure 3.19. I faintly detected EGFP at a higher molecular weight corresponding to either a fusion construct in my CMV-Myc-NKAP-P2A-EGFP sample, or resulting from NKAP displaying a higher apparent molecular weight due to being highly phosphorylated. I also faintly detected EGFP in my original CMV-EGFP-NKAP fusion construct that I created before redesigning my last two plasmids, although it was very weak.

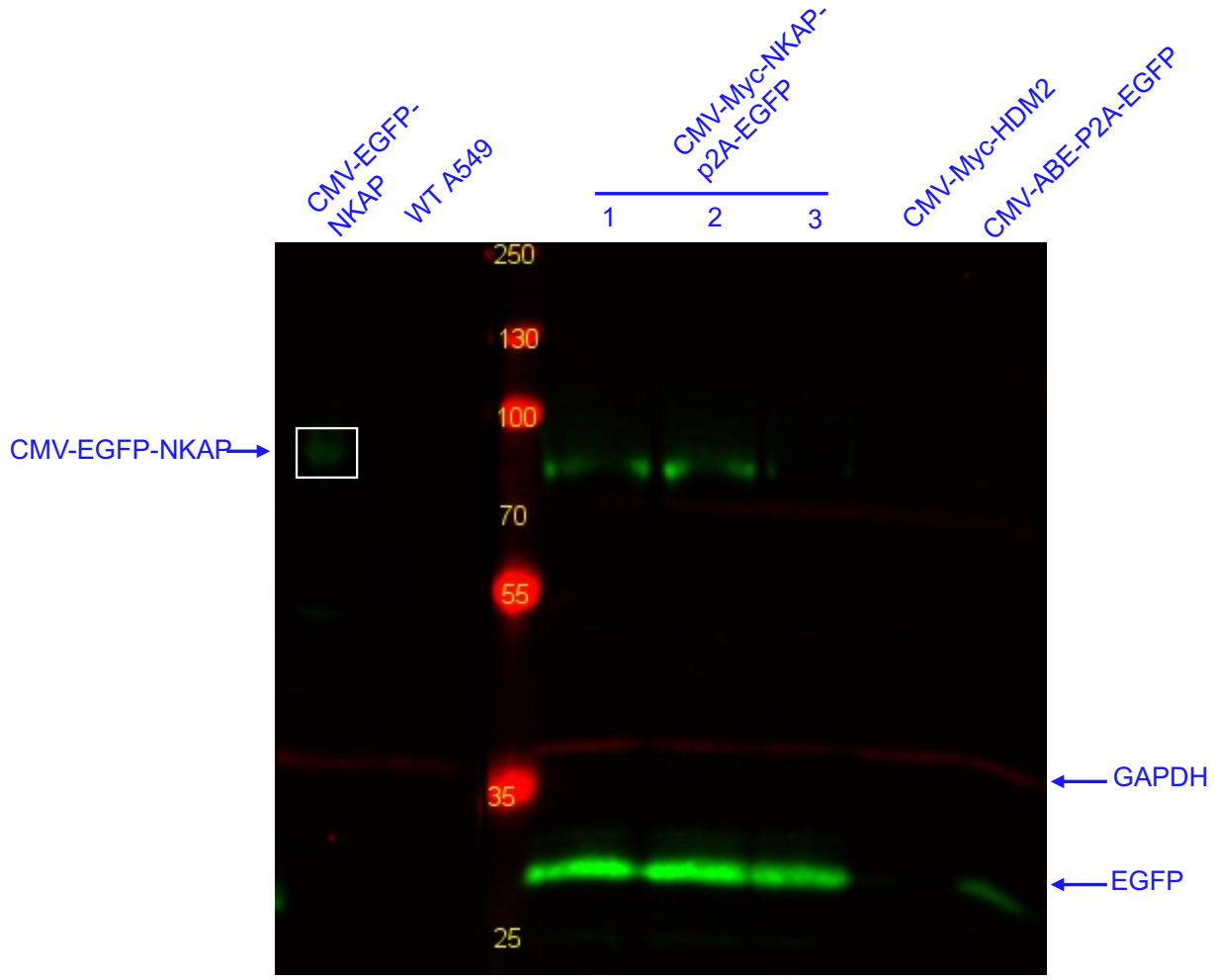


Figure 3.19 Western blot of A549 cell lysate following a transfection of CMV-EGFP-NKAP, CMV-Myc-NKAP-P2A-EGFP, CMV-Myc-HDM2, or CMV-ABE-P2A-EGFP. The blot was probed for EGFP on both pieces. Detection for GAPDH is very low due to the antibody being reused.

Table 3 Summary of the overexpression plasmids I created and their important features.

Plasmid	Description	Note
pNC1	CMV_NKAP	No Kozak consensus sequence; pcDNA3.2 backbone; HSV-poly(A)
pNC2	CMV_3xMyc-NKAP	
pNC3	CMV_GFP-linker-NKAP	
pNC4	CMV_GFP	
pNC5	CMV_GFP-linker-NKAP	Kozak sequence; bGHR-poly(A)
pNC6	CMV_3xMyc-NKAP_P2A-GFP	

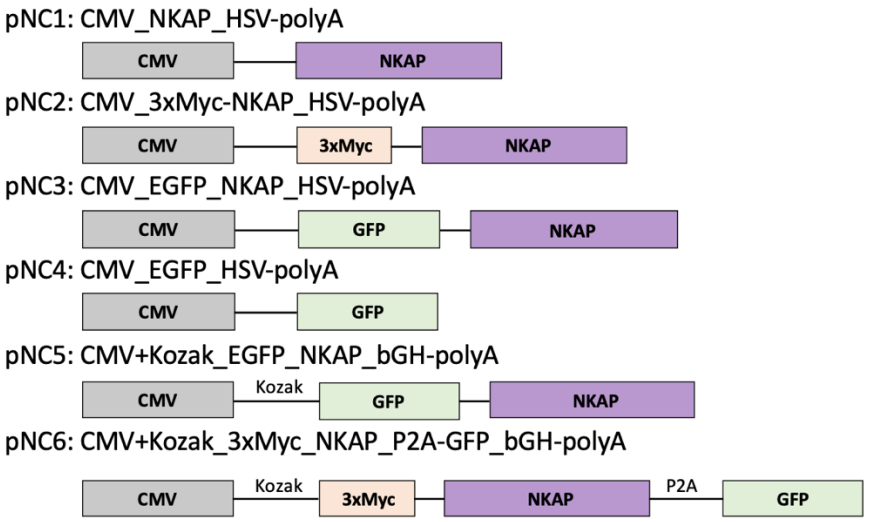


Figure 3.20 Final NKAP overexpression constructs.

Conclusions

The Kozak sequence is very important for expressing proteins from transiently transfected overexpression plasmids, but should not be present when expressing lncRNAs from overexpression plasmids with otherwise identical backbones. It appears that a small amount of NKAP mRNA was able to be loaded into the ribosome and translated, but not nearly at levels sufficient for overexpression.

In Figure 3.12, the microscope image for cells transfected with the control plasmid named pcEGFP (top middle) appears to fluoresce only sparsely. In brainstorming the optimization of my overexpression constructs, I considered if a random mutation in the chromophore of the EGFP that I had amplified out to clone my first four constructs could be the cause of the faint fluorescence. I aligned the plasmid pcEGFP to one of the brightly fluorescing controls and found that the green fluorescent proteins did not completely match. Upon further investigation, I found that the EGFP in the pcEGFP plasmid was in fact the original avGFP from Aequorea Victoria.[35] I suspect that the low detected of fluorescence is due to avGFP having a different excitation wavelength compared to EGFP. Due to various mutations that have optimized the fluorescence, EGFP has a peak excitation wavelength of 488 nm, whereas avGFP has a peak excitation wavelength of 508 nm. At 488 nm, only 23% of avGFP is excited.[36] The bright fluorescence

detected from EGFP leads me to suspect that our microscope's lasers are attuned to detect the shifted excitation wavelength of EGFP, but we would need a different laser to robustly detect avGFP.

In Figure 3.18, the CMV-Myc-NKAP-P2A-EGFP construct appears to either show an EGFP-NKAP fusion product, or NKAP running at a higher molecular weight. Indeed, the P2A sequence has been shown to cause fusions, although at levels less than 10%.[34] Additionally, NKAP has been detected at a higher molecular weight of 60-70+ kDa, as well as at its theoretical molecular weight of 47 kDa, depending on the sample and level of phosphorylation.[37,38] This is further veiled by photobleaching, instead of stripping, the western blot, as the primary antibodies for Myc and EGFP, respectively, are grown in the same animal. The secondary, infrared antibody (shown in green) detects the constant region of the primary antibodies, which is specific to the animal in which the antibody was generated. The resulting "signal crosstalk" helps to obscure the true identity of the band. However, based on the literature, it is my view that the band is primarily Myc-NKAP and not a Myc-NKAP-Fusion construct. My committee supports this hypothesis. Excising the band and performing mass spectrometry on its contents would be the only definite way to confirm the identity of the protein for this western blot. Additionally, repeating the western blot, cutting it, and probing different lanes with separate incubations of antibodies (rather than blotting sequentially) would aid in clarifying to what extent, if any, the fusion product is present.

References

(1) Park, E. G.; Ha, H.; Lee, D. H.; Kim, W. R.; Lee, Y. J.; Bae, W. H.; Kim, H.-S. Genomic Analyses of Non-Coding RNAs Overlapping Transposable Elements and Its Implication to Human Diseases. *IJMS* **2022**, *23* (16), 8950. https://doi.org/10.3390/ijms23168950.

(2) The ENCODE Project Consortium; Abascal, F.; Acosta, R.; Addleman, N. J.; Adrian, J.; Afzal, V.; Aken, B.; Ai, R.; Akiyama, J. A.; Jammal, O. A.; Amrhein, H.; Anderson, S. M.; Andrews, G. R.; Antoshechkin, I.; Ardlie, K. G.; Armstrong, J.; Astley, M.; Banerjee, B.; Barkal, A. A.; Barnes, I. H. A.; Barozzi, I.; Barrell, D.; Barson, G.; Bates, D.; Baymuradov, U. K.; Bazile, C.; Beer, M. A.; Beik, S.; Bender, M. A.; Bennett, R.; Bouvrette, L. P. B.; Bernstein, B. E.; Berry, A.; Bhaskar, A.; Bignell, A.; Blue, S. M.; Bodine, D. M.; Boix, C.; Boley, N.; Borrman, T.; Borsari, B.; Boyle, A. P.; Brandsmeier, L. A.; Breschi, A.; Bresnick, E. H.; Brooks, J. A.; Buckley, M.; Burge, C. B.; Byron, R.; Cahill, E.; Cai, L.; Cao, L.; Carty, M.; Castanon, R. G.; Castillo, A.; Chaib, H.; Chan, E. T.; Chee, D. R.; Chee, S.; Chen, H.; Chen, H.; Chen, J.-Y.; Chen, S.; Cherry, J. M.; Chhetri, S. B.; Choudhary, J. S.; Chrast, J.; Chung, D.; Clarke, D.; Cody, N. A. L.; Coppola, C. J.; Coursen, J.; D'Ippolito, A. M.; Dalton, S.; Danyko, C.; Davidson, C.; Davila-Velderrain, J.; Davis, C. A.; Dekker, J.; Deran, A.; DeSalvo, G.; Despacio-Reyes, G.; Dewey, C. N.; Dickel, D. E.; Diegel, M.; Diekhans, M.; Dileep, V.; Ding, B.; Djebali, S.; Dobin, A.; Dominguez, D.; Donaldson, S.; Drenkow, J.; Dreszer, T. R.; Drier, Y.; Duff, M. O.; Dunn, D.; Eastman, C.; Ecker, J. R.; Edwards, M. D.; El-Ali, N.; Elhajjajy, S. I.; Elkins, K.; Emili, A.; Epstein, C. B.; Evans, R. C.; Ezkurdia, I.; Fan, K.; Farnham, P. J.; Farrell, N.; Feingold, E. A.; Ferreira, A.-M.; Fisher-Aylor, K.; Fitzgerald, S.; Flicek, P.; Foo, C. S.; Fortier, K.; Frankish, A.; Freese, P.; Fu, S.; Fu, X.-D.; Fu, Y.; Fukuda-Yuzawa, Y.; Fulciniti, M.; Funnell, A. P. W.; Gabdank, I.; Galeev, T.; Gao, M.; Giron, C. G.; Garvin, T. H.; Gelboin-Burkhart, C. A.; Georgolopoulos, G.; Gerstein, M. B.; Giardine, B. M.; Gifford, D. K.; Gilbert, D. M.; Gilchrist, D. A.; Gillespie, S.; Gingeras, T. R.; Gong, P.; Gonzalez, A.; Gonzalez, J. M.; Good, P.; Goren, A.; Gorkin, D. U.; Graveley, B. R.; Gray, M.; Greenblatt, J. F.; Griffiths, E.; Groudine, M. T.; Grubert, F.; Gu, M.; Guigó, R.; Guo, H.; Guo, Y.; Guo, Y.; Gursoy, G.; Gutierrez-Arcelus, M.; Halow, J.; Hardison, R. C.; Hardy, M.; Hariharan, M.; Harmanci, A.; Harrington, A.; Harrow, J. L.; Hashimoto, T. B.; Hasz, R. D.; Hatan, M.; Haugen, E.; Hayes, J. E.; He, P.; He, Y.; Heidari, N.; Hendrickson, D.; Heuston, E. F.; Hilton, J. A.; Hitz, B. C.; Hochman, A.; Holgren, C.; Hou, L.; Hou, S.; Hsiao, Y.-H. E.; Hsu, S.; Huang, H.; Hubbard, T. J.; Huey, J.; Hughes, T. R.; Hunt, T.; Ibarrientos, S.; Issner, R.; Iwata, M.; Izuogu, O.; Jaakkola, T.; Jameel, N.; Jansen, C.; Jiang, L.; Jiang, P.; Johnson, A.; Johnson, R.; Jungreis, I.; Kadaba, M.; Kasowski, M.; Kasparian, M.; Kato, M.; Kaul, R.; Kawli, T.; Kay, M.; Keen, J. C.; Keles, S.; Keller, C. A.; Kelley, D.; Kellis, M.; Kheradpour, P.; Kim, D. S.; Kirilusha, A.; Klein, R. J.; Knoechel, B.; Kuan, S.; Kulik, M. J.; Kumar, S.; Kundaje, A.; Kutyavin, T.; Lagarde, J.; Lajoie, B. R.; Lambert, N. J.; Lazar, J.; Lee, A. Y.; Lee, D.; Lee, E.; Lee, J. W.; Lee, K.; Leslie, C. S.; Levy, S.; Li, B.; Li, H.; Li, N.; Li, S.; Li, X.; Li, Y. I.; Li, Y.; Li, Y.; Li, Y.; Lian, J.; Libbrecht, M. W.; Lin, S.; Lin, Y.; Liu, D.; Liu, J.; Liu, P.; Liu, T.; Liu, X. S.; Liu, Y.; Liu, Y.; Long, M.; Lou, S.; Loveland, J.; Lu, A.; Lu, Y.; Lécuyer, E.; Ma, L.; Mackiewicz, M.; Mannion, B. J.; Mannstadt, M.; Manthravadi, D.; Marinov, G. K.; Martin, F. J.; Mattei, E.; McCue, K.; McEown, M.; McVicker, G.; Meadows, S. K.; Meissner, A.; Mendenhall, E. M.; Messer, C. L.; Meuleman, W.; Meyer, C.; Miller, S.; Milton, M. G.; Mishra, T.; Moore, D. E.; Moore, H. M.; Moore, J. E.; Moore, S. H.; Moran, J.; Mortazavi, A.; Mudge, J. M.; Munshi, N.; Murad, R.; Myers, R. M.; Nandakumar, V.; Nandi, P.; Narasimha, A. M.; Narayanan, A. K.; Naughton, H.; Navarro, F. C. P.; Navas, P.; Nazarovs, J.; Nelson, J.; Neph, S.; Neri, F. J.; Nery, J. R.; Nesmith, A. R.; Newberry, J. S.; Newberry, K. M.; Ngo, V.; Nguyen, R.; Nguyen, T. B.; Nguyen, T.; Nishida, A.; Noble, W. S.; Novak, C. S.; Novoa, E. M.; Nuñez, B.; O'Donnell, C. W.; Olson, S.; Onate, K. C.; Otterman, E.; Ozadam, H.; Pagan,

M.; Palden, T.; Pan, X.; Park, Y.; Partridge, E. C.; Paten, B.; Pauli-Behn, F.; Pazin, M. J.; Pei, B.; Pennacchio, L. A.; Perez, A. R.; Perry, E. H.; Pervouchine, D. D.; Phalke, N. N.; Pham, Q.; Phanstiel, D. H.; Plajzer-Frick, I.; Pratt, G. A.; Pratt, H. E.; Preissl, S.; Pritchard, J. K.; Pritykin, Y.; Purcaro, M. J.; Qin, Q.; Quinones-Valdez, G.; Rabano, I.; Radovani, E.; Raj, A.; Rajagopal, N.; Ram, O.; Ramirez, L.; Ramirez, R. N.; Rausch, D.; Raychaudhuri, S.; Raymond, J.; Razavi, R.; Reddy, T. E.; Reimonn, T. M.; Ren, B.; Reymond, A.; Reynolds, A.; Rhie, S. K.; Rinn, J.; Rivera, M.; Rivera-Mulia, J. C.; Roberts, B.; Rodriguez, J. M.; Rozowsky, J.; Ryan, R.; Rynes, E.; Salins, D. N.; Sandstrom, R.; Sasaki, T.; Sathe, S.; Savic, D.; Scavelli, A.; Scheiman, J.; Schlaffner, C.; Schloss, J. A.; Schmitges, F. W.; See, L. H.; Sethi, A.; Setty, M.; Shafer, A.; Shan, S.; Sharon, E.; Shen, Q.; Shen, Y.; Sherwood, R. I.; Shi, M.; Shin, S.; Shoresh, N.; Siebenthall, K.; Sisu, C.; Slifer, T.; Sloan, C. A.; Smith, A.; Snetkova, V.; Snyder, M. P.; Spacek, D. V.; Srinivasan, S.; Srivas, R.; Stamatoyannopoulos, G.; Stamatoyannopoulos, J. A.; Stanton, R.; Steffan, D.; Stehling-Sun, S.; Strattan, J. S.; Su, A.; Sundararaman, B.; Suner, M.-M.; Syed, T.; Szynkarek, M.; Tanaka, F. Y.; Tenen, D.; Teng, M.; Thomas, J. A.; Toffey, D.; Tress, M. L.; Trout, D. E.; Trynka, G.; Tsuji, J.; Upchurch, S. A.; Ursu, O.; Uszczynska-Ratajczak, B.; Uziel, M. C.; Valencia, A.; Biber, B. V.; Van Der Velde, A. G.; Van Nostrand, E. L.; Vaydylevich, Y.; Vazquez, J.; Victorsen, A.; Vielmetter, J.; Vierstra, J.; Visel, A.; Vlasova, A.; Vockley, C. M.; Volpi, S.; Vong, S.; Wang, H.; Wang, M.; Wang, Q.; Wang, R.; Wang, T.; Wang, W.; Wang, X.; Wang, Y.; Watson, N. K.; Wei, X.; Wei, Z.; Weisser, H.; Weissman, S. M.; Welch, R.; Welikson, R. E.; Weng, Z.; Westra, H.-J.; Whitaker, J. W.; White, C.; White, K. P.; Wildberg, A.; Williams, B. A.; Wine, D.; Witt, H. N.; Wold, B.; Wolf, M.; Wright, J.; Xiao, R.; Xiao, X.; Xu, J.; Xu, J.; Yan, K.-K.; Yan, Y.; Yang, H.; Yang, X.; Yang, Y.-W.; Yardımcı, G. G.; Yee, B. A.; Yeo, G. W.; Young, T.; Yu, T.; Yue, F.; Zaleski, C.; Zang, C.; Zeng, H.; Zeng, W.; Zerbino, D. R.; Zhai, J.; Zhan, L.; Zhan, Y.; Zhang, B.; Zhang, J.; Zhang, J.; Zhang, K.; Zhang, L.; Zhang, P.; Zhang, Q.; Zhang, X.-O.; Zhang, Y.; Zhang, Z.; Zhao, Y.; Zheng, Y.; Zhong, G.; Zhou, X.-Q.; Zhu, Y.; Zimmerman, J.; Snyder, M. P.; Gingeras, T. R.; Moore, J. E.; Weng, Z.; Gerstein, M. B.; Ren, B.; Hardison, R. C.; Stamatoyannopoulos, J. A.; Graveley, B. R.; Feingold, E. A.; Pazin, M. J.; Pagan, M.; Gilchrist, D. A.; Hitz, B. C.; Cherry, J. M.; Bernstein, B. E.; Mendenhall, E. M.; Zerbino, D. R.; Frankish, A.; Flicek, P.; Myers, R. M. Perspectives on ENCODE. *Nature* **2020**, *583* (7818), 693–698. https://doi.org/10.1038/s41586-020-2449-8.

(3) Mercer, T. R.; Dinger, M. E.; Mattick, J. S. Long Non-Coding RNAs: Insights into Functions. *Nat Rev Genet* **2009**, *10* (3), 155–159. https://doi.org/10.1038/nrg2521.

(4) Mattick, J. S.; Amaral, P. P.; Carninci, P.; Carpenter, S.; Chang, H. Y.; Chen, L.-L.; Chen, R.; Dean, C.; Dinger, M. E.; Fitzgerald, K. A.; Gingeras, T. R.; Guttman, M.; Hirose, T.; Huarte, M.; Johnson, R.; Kanduri, C.; Kapranov, P.; Lawrence, J. B.; Lee, J. T.; Mendell, J. T.; Mercer, T. R.; Moore, K. J.; Nakagawa, S.; Rinn, J. L.; Spector, D. L.; Ulitsky, I.; Wan, Y.; Wilusz, J. E.; Wu, M. Long Non-Coding RNAs: Definitions, Functions, Challenges and Recommendations. *Nat Rev Mol Cell Biol* **2023**, *24* (6), 430–447. https://doi.org/10.1038/s41580-022-00566-8.

(5) Wilusz, J. E.; Freier, S. M.; Spector, D. L. 3′ End Processing of a Long Nuclear-Retained Noncoding RNA Yields a tRNA-like Cytoplasmic RNA. *Cell* **2008**, *135* (5), 919–932. https://doi.org/10.1016/j.cell.2008.10.012.

(6) Gingeras, T. R. Origin of Phenotypes: Genes and Transcripts. *Genome Res.* **2007**, *17* (6), 682–690. https://doi.org/10.1101/gr.6525007.

(7) Kazemzadeh, M.; Safaralizadeh, R.; Orang, A. V. LncRNAs: Emerging Players in Gene Regulation and Disease Pathogenesis. *J Genet* **2015**, *94* (4), 771–784. https://doi.org/10.1007/s12041-015-0561-6.

(8) Parasramka, M. A.; Maji, S.; Matsuda, A.; Yan, I. K.; Patel, T. Long Non-Coding RNAs as Novel Targets for Therapy in Hepatocellular Carcinoma. *Pharmacology & Therapeutics* **2016**, *161*, 67–78. https://doi.org/10.1016/j.pharmthera.2016.03.004.

(9) Su, T.; Trang, N.; Zhu, J.; Kong, L.; Cheung, D.; Chou, V.; Ellis, L.; Huang, C.; Camden, N.; McHugh, C. A. GRAS1 Non-Coding RNA Protects against DNA Damage and Cell Death by Binding and Stabilizing NKAP. June 20, 2023. https://doi.org/10.1101/2023.06.20.545783.

(10) Gao, J.; Aksoy, B. A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S. O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; Cerami, E.; Sander, C.; Schultz, N. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci Signal* **2013**, *6* (269), pl1. https://doi.org/10.1126/scisignal.2004088.

(11) Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B. E.; Sumer, S. O.; Aksoy, B. A.; Jacobsen, A.; Byrne, C. J.; Heuer, M. L.; Larsson, E.; Antipin, Y.; Reva, B.; Goldberg, A. P.; Sander, C.; Schultz, N. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* **2012**, *2* (5), 401–404. https://doi.org/10.1158/2159-8290.CD-12-0095.

(12) Chen, D.; Li, Z.; Yang, Q.; Zhang, J.; Zhai, Z.; Shu, H.-B. Identification of a Nuclear Protein That Promotes NF-κB Activation. *Biochemical and Biophysical Research Communications* **2003**, *310* (3), 720–724. https://doi.org/10.1016/j.bbrc.2003.09.074.

(13) Pajerowski, A. G.; Nguyen, C.; Aghajanian, H.; Shapiro, M. J.; Shapiro, V. S. NKAP Is a Transcriptional Repressor of Notch Signaling and Is Required for T Cell Development. *Immunity* **2009**, *30* (5), 696–707. https://doi.org/10.1016/j.immuni.2009.02.011.

(14) Ma, Q.; Hou, L.; Gao, X.; Yan, K. NKAP Promotes Renal Cell Carcinoma Growth via AKT/mTOR Signalling Pathway. *Cell Biochem Funct* **2020**, *38* (5), 574–581. https://doi.org/10.1002/cbf.3508.

(15) Gu, G.; Gao, T.; Zhang, L.; Chen, X.; Pang, Q.; Wang, Y.; Wang, D.; Li, J.; Liu, Q. NKAP Alters Tumor Immune Microenvironment and Promotes Glioma Growth via Notch1 Signaling. *J Exp Clin Cancer Res* **2019**, *38* (1), 291. https://doi.org/10.1186/s13046-019-1281-1.

(16) 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**, *42* (5), 2130–2138. https://doi.org/10.3892/or.2019.7322.

(17) 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**, *13* (1), 73. https://doi.org/10.1038/s41419-022-04524-2.

(18) Zhang, J.; Bai, R.; Li, M.; Ye, H.; Wu, C.; Wang, C.; Li, S.; Tan, L.; Mai, D.; Li, G.; Pan, L.; Zheng, Y.; Su, J.; Ye, Y.; Fu, Z.; Zheng, S.; Zuo, Z.; Liu, Z.; Zhao, Q.; Che, X.; Xie, D.; Jia,

W.; Zeng, M.-S.; Tan, W.; Chen, R.; Xu, R.-H.; Zheng, J.; Lin, D. Excessive miR-25-3p Maturation via N6-Methyladenosine Stimulated by Cigarette Smoke Promotes Pancreatic Cancer Progression. *Nat Commun* **2019**, *10* (1), 1858. https://doi.org/10.1038/s41467-019-09712-x.

(19) Fiordaliso, S. K.; Iwata-Otsubo, A.; Ritter, A. L.; Quesnel-Vallières, M.; Fujiki, K.; Nishi, E.; Hancarova, M.; Miyake, N.; Morton, J. E. V.; Lee, S.; Hackmann, K.; Bando, M.; Masuda, K.; Nakato, R.; Arakawa, M.; Bhoj, E.; Li, D.; Hakonarson, H.; Takeda, R.; Harr, M.; Keena, B.; Zackai, E. H.; Okamoto, N.; Mizuno, S.; Ko, J. M.; Valachova, A.; Prchalova, D.; Vlckova, M.; Pippucci, T.; Seiler, C.; Choi, M.; Matsumoto, N.; Di Donato, N.; Barash, Y.; Sedlacek, Z.; Shirahige, K.; Izumi, K. Missense Mutations in NKAP Cause a Disorder of Transcriptional Regulation Characterized by Marfanoid Habitus and Cognitive Impairment. *The American Journal of Human Genetics* **2019**, *105* (5), 987–995. https://doi.org/10.1016/j.ajhg.2019.09.009.

(20) Xu, X.; Gao, C.; Ye, F.; Peng, A.; Xu, J.; Jin, K.; Zhang, J.; Ye, Y.; Yang, Y.; Zhang, X.; Shen, S.; Jin, F. From Phenotype to Mechanism: Prenatal Spectrum of *NKAP* Mutation-related Disorder and Its Pathogenesis Inducing Congenital Heart Disease. *J Cellular Molecular Medi* **2024**, *28* (8), e18305. https://doi.org/10.1111/jcmm.18305.

(21) 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.; Mandegar, M. A.; Olvera, M. P.; Gilbert, L. A.; Conklin, B. R.; Chang, H. Y.; Weissman, J. S.; Lim, D. A. CRISPRi-Based Genome-Scale Identification of Functional Long Noncoding RNA Loci in Human Cells. *Science* **2017**, *355* (6320), eaah7111. https://doi.org/10.1126/science.aah7111.

(22) Lim, K. R. Q.; Yokota, T. Invention and Early History of Gapmers. In *Gapmers*; Yokota, T., Maruyama, R., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2020; Vol. 2176, pp 3–19. https://doi.org/10.1007/978-1-0716-0771-8_1.

(23) McHugh, C. A.; Guttman, M. RAP-MS: A Method to Identify Proteins That Interact Directly with a Specific RNA Molecule in Cells. In *RNA Detection*; Gaspar, I., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2018; Vol. 1649, pp 473–488. https://doi.org/10.1007/978-1-4939-7213-5_31.

(24) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29* (1), 15–21. https://doi.org/10.1093/bioinformatics/bts635.

(25) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25* (16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

(26) Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y. C.; Laslo, P.; Cheng, J. X.; Murre, C.; Singh, H.; Glass, C. K. Simple Combinations of Lineage-Determining Transcription Factors Prime Cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* **2010**, *38* (4), 576–589. https://doi.org/10.1016/j.molcel.2010.05.004.

(27) Dixon, S. J.; Lemberg, K. M.; Lamprecht, M. R.; Skouta, R.; Zaitsev, E. M.; Gleason, C. E.; Patel, D. N.; Bauer, A. J.; Cantley, A. M.; Yang, W. S.; Morrison, B.; Stockwell, B. R.

Ferroptosis: An Iron-Dependent Form of Nonapoptotic Cell Death. *Cell* **2012**, *149* (5), 1060–1072. https://doi.org/10.1016/j.cell.2012.03.042.

(28) Urdaneta, E. C.; Vieira-Vieira, C. H.; Hick, T.; Wessels, H.-H.; Figini, D.; Moschall, R.; Medenbach, J.; Ohler, U.; Granneman, S.; Selbach, M.; Beckmann, B. M. Purification of Cross-Linked RNA-Protein Complexes by Phenol-Toluol Extraction. *Nat Commun* **2019**, *10* (1), 990. https://doi.org/10.1038/s41467-019-08942-3.

(29) Van Nostrand, E. L.; Pratt, G. A.; Shishkin, A. A.; Gelboin-Burkhart, C.; Fang, M. Y.; Sundararaman, B.; Blue, S. M.; Nguyen, T. B.; Surka, C.; Elkins, K.; Stanton, R.; Rigo, F.; Guttman, M.; Yeo, G. W. Robust Transcriptome-Wide Discovery of RNA-Binding Protein Binding Sites with Enhanced CLIP (eCLIP). *Nat Methods* **2016**, *13* (6), 508–514. https://doi.org/10.1038/nmeth.3810.

(30) Spitzer, N.; Sammons, G. S.; Price, E. M. Autofluorescent Cells in Rat Brain Can Be Convincing Impostors in Green Fluorescent Reporter Studies. *Journal of Neuroscience Methods* **2011**, *197* (1), 48–55. https://doi.org/10.1016/j.jneumeth.2011.01.029.

(31) Kozak, M. At Least Six Nucleotides Preceding the AUG Initiator Codon Enhance Translation in Mammalian Cells. *Journal of Molecular Biology* **1987**, *196* (4), 947–950. https://doi.org/10.1016/0022-2836(87)90418-9.

(32) Lesueur, L. L.; Mir, L. M.; André, F. M. Overcoming the Specific Toxicity of Large Plasmids Electrotransfer in Primary Cells In Vitro. *Molecular Therapy - Nucleic Acids* **2016**, *5*, e291. https://doi.org/10.1038/mtna.2016.4.

(33) Cowan, Q. T.; Gu, S.; Gu, W.; Ranzau, B. L.; Simonson, T. S.; Komor, A. C. Development of Multiplexed Orthogonal Base Editor (MOBE) Systems. *Nat Biotechnol* **2024**. https://doi.org/10.1038/s41587-024-02240-0.

(34) Liu, Z.; Chen, O.; Wall, J. B. J.; Zheng, M.; Zhou, Y.; Wang, L.; Ruth Vaseghi, H.; Qian, L.; Liu, J. Systematic Comparison of 2A Peptides for Cloning Multi-Genes in a Polycistronic Vector. *Sci Rep* **2017**, *7* (1), 2193. https://doi.org/10.1038/s41598-017-02460-2.

(35) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary Structure of the Aequorea Victoria Green-Fluorescent Protein. *Gene* **1992**, *111* (2), 229–233. https://doi.org/10.1016/0378-1119(92)90691-H.

(36) Lambert, T. J. FPbase: A Community-Editable Fluorescent Protein Database. *Nat Methods* **2019**, *16* (4), 277–278. https://doi.org/10.1038/s41592-019-0352-8.

(37) Li, T.; Chen, L.; Cheng, J.; Dai, J.; Huang, Y.; Zhang, J.; Liu, Z.; Li, A.; Li, N.; Wang, H.; Yin, X.; He, K.; Yu, M.; Zhou, T.; Zhang, X.; Xia, Q. SUMOylated NKAP Is Essential for Chromosome Alignment by Anchoring CENP-E to Kinetochores. *Nat Commun* **2016**, *7* (1), 12969. https://doi.org/10.1038/ncomms12969.

(38) Burgute, B. D.; Peche, V. S.; Steckelberg, A.-L.; Glöckner, G.; Gaßen, B.; Gehring, N. H.; Noegel, A. A. NKAP Is a Novel RS-Related Protein That Interacts with RNA and RNA Binding Proteins. *Nucleic Acids Research* **2014**, *42* (5), 3177–3193. https://doi.org/10.1093/nar/gkt1311.