Spring 5-2019

An Assessment of Combinatorial Transcription Factor Activity at p53 Enhancer Elements

Sylvia Kuang
University at Albany, State University of New York, skuang@albany.edu

An assessment of combinatorial transcription factor activity at p53 family enhancer elements

An honors thesis presented to the
Department of Biological Sciences
University at Albany, State University of New York
in partial fulfillment of the requirements
for graduation with Honors in Biological Sciences
and
graduation from The Honors College.

Sylvia Kuang

Research Mentor: Morgan Sammons, Ph.D.

May 2019

Abstract

Certain non-coding DNA sequences in the eukaryotic genome regulate gene expression. These non-coding regulatory regions, including promoters and enhancers, are controlled by the binding of multiple transcription factors which act together to regulate gene transcription. The number of potential transcription factor combinations regulating any gene presents a massive experimental challenge. One well-known transcription factor, p53, activates multiple transcription pathways involved in tumor suppression, primarily through engagement with enhancers. p53 is one member of a paralogous transcription factor family, which includes the factor p63. Whereas p53 is involved in tumor suppression, p63 is a transcription factor responsible for maintaining epithelial cell populations through its ability to bind to and regulate enhancers. p63 and p53 are often bound to the same enhancers in the genome, suggesting a more complex regulation than predicted by their canonical functions. We therefore aimed to better understand how genomic binding sequences and other factors regulate p53 and p63 activity at enhancers. Luciferase reporter gene assays were utilized to measure the transcriptional output of various p63 and p53 enhancers after genetically altering flanking DNA sequence motifs. We found that changing these flanking regions revealed core regulatory sequences that drive p53 and p63 transcriptional activity. We also determined that p63-bound enhancers, but not those bound by p53, had context-dependent activity. Depending on the cell type, these enhancers are active or inactive, with basal expression of p63 determining their activity. Our data provide new insight into the regulation of p53 family enhancers, and further work will lead to a better understanding of transcription factor activity and function at enhancers.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Morgan Sammons for the continuous support of my research. Without his patience, guidance, and immense knowledge, I would not have been able to reach my goals. I could not have imagined having a better advisor and mentor for my undergraduate research experience.

In addition to my advisor, I would like to sincerely thank Ally Catizone, the graduate student I worked directly under. She provided the direction for all of my work, in addition to endless hours helping me author papers and interpret data. I owe much of my success to her endless patience and encouragement while being my mentor.

My sincere thanks also goes to my other fellow lab members: Serene Durham, Dr. Gizem Uzunbas, Chelsi Riley, and Kate Sazon, who all contributed to making the Sammons Lab a wonderful and nurturing academic environment.

List of Figures

Table of Contents

Introduction

Genes are sequences of DNA which can vary from a few hundred to more than 2 million base pairs in length (Biscotnini, 2019). The information within genes can be decoded by a process called transcription, whereby the DNA information is transcribed into a new molecule called mRNA (messenger RNA). The information on an mRNA is then decoded by ribosomes to produce a polypeptide. Gene transcription is tightly controlled by gene regulatory networks to ensure the correct spatial and temporal expression of the gene. This regulation occurs through regulatory elements, which include enhancers and promoters. Promoters are DNA sequences upstream of a target gene, where transcription factors and RNA polymerase bind to a transcription start site in order to initiate transcription. While promoters are strictly required for transcription, enhancers are not required for basal transcription. However, enhancers control the abundance and frequency of transcription, as well as the spatial and temporal expression of genes in the larger organism. Both regulatory elements function through the combinatorial binding of transcription factors, which direct and regulate the general transcription machinery. The expression of different sets of genes in different tissues and at different times is known as differential gene expression. This requires the coordinated effort of multiple transcription factors, which then binds to different combinations of enhancers and promoters.

Enhancers are unique, non-coding sections of the eukaryotic genome that regulate transcription by recruiting transcription factors. Enhancers are a few hundred base pairs in length, and they can be located at various distances from the target gene, up to a mega base away. However, transcription factors rarely act alone at enhancers. The binding of the correct combination and orientation of factors to both the enhancer and promoter regions initiates

transcription of the gene. Their activity should therefore be studied in combination with other transcription factors (Spitz & Furlong, 2012).

The p53 family of transcription factors includes p53, p63, and p73, and is highly enriched at enhancer regions. These proteins can bind to the same DNA sites, have similar DNA binding domains, and can activate some of the same target genes. However, there is evidence that they are not identical nor redundant, as previous studies have demonstrated that the loss of one family member during embryonic mouse development results in severe deformations, if not lethality (Nostrand et al., 2017). While p53 has a major role in tumor suppression, p63 is responsible for epithelial morphogenesis, and p73 is known to maintain neuronal structures. We are primarily focused on the activities and functions of p53 and p63. p53 is a transcription factor that is expressed in every cell type. It acts a trans-activator, which enhances the transcription of p53 target genes by recruiting histone-modifiers (Brady & Attardi, 2010). Furthermore, p53 is mutated in 50% of cancers due to its crucial role in the regulation of downstream genes responsible for cell cycle arrest, senescence, autophagy, and apoptosis (Zilfou & Lowe, 2009).

While p63 shares some of the same target genes with p53, p63 primarily regulates downstream targets that drive the early development of epidermal structures and their functions (Koster, 2010). Loss of both proteins is known to compromise cell senescence and apoptosis, leading to many different cancers and diseases. Previous studies have shown that p63-knockout mice develop with truncated limbs, the absence of a stratified epidermis, and deformations of other appendages (Romano et al., 2012). The literature also suggests that p63 can and does interact with the p53 network. However, it is unknown what defines a p63 enhancer, what the role of p63 is at enhancers, and what the target genes of p63-dependent enhancers are.

A previous experiment in our lab screened hundreds of p53 and p63 enhancers including a single group which all bind p63, but do not bind p53. We found this interesting because both proteins bind a highly similar DNA response element and often bind the same regions. However, this group of p63 enhancers showed minimal activity in a massively parallel reporter assay (MPRA) when the p63 protein was present and did not respond to p53 induction. Loss of p63 or p63 binding site led to decreased enhancer activity. This study also shows that the loss of p53 and/or p63 had detrimental effects on p53/p63 bound enhancers but positive effects on p53/p63 independent enhancers. We wanted to identify the gene associated with these seemingly p63 bound but lowly active enhancers *in vivo.* We hope to dissect these regions and to understand the interplay between p53 and p63 using these locations. In order to further investigate these results, we chose a number of enhancers from this p63-only group. Luciferase assays were performed using two different cell lines in order to study the effect of the presence and absence of p63 in the native cell environment on enhancer activity. The enhancers were chosen based on their location to genes that were downregulated in p63 knockdown cell lines, their enrichment of p63 protein, and the presence of active enhancer chromatin marks.

To study the activity of these enhancers, dCas9-KRAB was targeted to the p63 binding sites of each of the enhancers. RNA sequencing was utilized to measure the output of these ten blocked enhancers to determine their enhancer functions. Further experiments can then be done to identify the core regulatory sequences that determine the activity of these enhancers. Ultimately, better knowledge of how p53 and p63 enhancers function under stress conditions will lead to a better understanding of their full impact on gene expression.

Materials and Methods

I. Cloning cluster 2 enhancers for Luciferase assays

1. Construction of reporter plasmid

The three enhancers were first given a numeric code: E231 for CUX1, E245 for CTSV, and E220 for DTHD1. Primers were designed using IDT PrimerQuest Tool to amplify enhancer regions from human genomic DNA. SnapGene and an online primer design tool (NEBuilder) were used to generate primers for Gibson assembly into pGL4.24, and were purchased from IDT. The enhancer sizes were verified by running them on a 2% agarose gel with a 100bp ladder. The three bands containing the different enhancer sequences were then cut from the gel for DNA extraction and purification.

Additionally, the pGL4.24 plasmid was restriction digested using the HindIII and KpnI restriction enzyme sites. The size of the cut plasmid was then verified by running a 1% agarose gel with a 1000bp ladder. The band was excised from the gel and purified using the NEB Monarch Gel Extraction kit. Finally, Gibson assembly was used to ligate each enhancer to a pGL4.24 plasmid at the HindIII and KpnI sites using the NEB HiFi Genome Assembly kit. Ligation products were selected after transformation and growth on LB-Ampicillin plates, and DNA sequencing was performed to verify the correct sequence and orientation of the ligated plasmid.

2. Bacterial transformation

The reporter assay plasmids were transformed into chemically-competent bacterial cells (*stbl3* from New England BioLabs) and then plated onto LB-Ampicillin plates. One colony per

enhancer was picked and grown overnight in 5mL of liquid culture at 37˚C in shaking culture. Plasmid DNA was then isolated using standard silica column and alkaline lysis methodologies.

3. Luciferase assays of cluster 2 enhancers in HCT116 and MCF10a cells

HCT116 WT, HCT116 ΔNp63, MCF10a WT, and MCF10a p63KD cells were seeded at 150,000 cells into 6-well dishes. The cell lines were then transfected with either one of the cluster 2 enhancers (E220, E231, E245), a no-enhancer negative control, or an ATF3 positive control using JetPrime transfection reagent. Dual Luciferase assays (Promega Dual-Glo) were then performed on each of the cell lines in three technical and three biological replicates as per manufacturer's standard recommendations.

II. Cloning cluster 2 enhancers for *in vivo* analysis of p63 dependent enhancer activity

1. Construction of reporter plasmid

The ten randomly selected Cluster 2 enhancers were first given a numeric code: 197, 198, 200, 206, 213, 236, 254, 258, 275, 288 in order to blind us to their identity during future experiments. One to two guide RNAs per enhancer were designed using the MIT CRISPR Design (reference) tool to target each of the enhancer sequences and were purchased from IDT (Integrated DNA Technologies). Additionally, the parental KRAB plasmid (Addgene plasmid #110820, dCas9-KRAB was a gift from Alejandro Chavez & George Church) was restriction digested using the Esp31 (BsmBI) restriction enzyme sites and then 5' phosphates were removed using shrimp alkaline phosphatase (SAP). Primer pairs were phosphorylated at the 5' end using T4 PNK and ATP at 37˚C for 30 minutes. Phosphorylated primers were then annealed by mixing equimolar amounts (10mM) in water, heating to 95˚C, followed by controlled cooling to 25˚C (at

a rate of 1˚C/minute) in a thermocycler. Annealed primers were diluted to 10uM and used in a ligation reaction with dephosphorylated backbone and T4 DNA ligase.

2. Bacterial transformation

Ligation products were transformed into New England BioLabs *stbl3 E.coli* cells and then plated onto LB-Ampicillin plates for selection of correctly ligated products. Plasmids were isolated from bacterial transformations and sent for Sanger sequencing to test for proper ligation products.

3. Viral transduction

HEK293FT cells were seeded at 900,000 cells per well of a 6-well dish. The cell lines were then transfected with sequence-verified lentiviral-dCas9-KRAB-gRNA constructs targeting either one of the 10 test enhancers, a known enhancer positive control (GDF15), or the negative control (FGF2 enhancer, inactive in MCF10A cells). The viral media was collected after 24 hours and again after 48 hours, filtered, and 8μg/μL of polybrene was added to increase transduction efficiency. MCF10A cells were seeded at 150,000 cells into 6-well dishes. The cell lines were then transduced with 1mL each of the twelve viral preparations, along with a negative control containing no virus. Cells were selected 48 hours post-transduction with 2μg/mL Puromycin for 24 hours. Cells surviving selection were cryopreserved or used directly in experiments.

4. RNA-seq analysis

Total RNA was extracted and purified from virally-transduced cell lines using the EZNA Total RNA Kit (Omega BioTek). An RNA-seq library was made using the BioO NEXTflex Rapid Directional RNA-Seq method. To summarize, 1ug of total RNA was used as the starting point for polyA-RNA selection using poly-dT-coupled magnetic beads. The resulting polyA-RNA was fragmented to an average size of 250bp. These fragments were then used for randomly primed 1st strand synthesis, followed by 2nd strand synthesis using RNAseH-mediated nicking and substitution of dTTP with dUTP. This allows for downstream detection of the original strand, thus allowing quantification of strandedness. Sequencing adapters and PCR were then performed to add both required sequencer information and unique barcode sequences. The quality of the resulting library was confirmed using a Qubit fluorimeter, qPCR (using the BioRad iTaq Universal SYBR Green One-Step Kit), and with capillary electrophoresis (Agilent Bioanalyzer at the University at Albany Cancer Research Center). All 12 libraries were pooled at an equal molar amount and a 1x75bp sequencing run was performed using a NextSeq 500 instrument. Transcript abundance was determined using *salmon (*reference) against the hg19 RefSeq genome assembly.

5. MIR205 CRISPR-KRAB-dCas9 qPCR

MCF10a WT cells were seeded at ??? cells into a 6-well dish. The cell lines were then transfected with either E197, the positive control (GDF15), or no virus. The cells were then selected 48 hours post-transduction with 2μg/mL Puromycin for 24 hours. Transduced cells were lysed and the total RNA was extracted and purified using the ENZA Total RNA Kit. qPCR was then performed on cDNA generated from total extracted RNA.

Results

We began by randomly selecting three enhancer sequences from the previously performed MPRA experiment. The three p63 enhancer sequences (E245, E231, E220) were first amplified out of genomic DNA utilizing PCR. The primers were designed to encompass what we believed represented the entire enhancer sequence, using DNAseI-hypersensitivity data from the ENCODE Project (Thurman et al., 2012). The primers were designed to have 5' homology to the enhancer test vector pGL4.24 in order to allow for downstream Gibson cloning. The enhancer sizes ranged from 265 to 530bp after considering the entirety of the DNAseI-hypersensitive region, as confirmed by 2% agarose gel analysis with a 100bp ladder. The three enhancer sequences amplified out of genomic DNA match with their expected amplicon sizes based on their primer sequences (Fig.1). The bands containing the different enhancer sequences were then cut from the gel for DNA extraction and purification and the resulting concentrations were measured with the Qubit High Sensitivity dsDNA kit and the Qubit fluorimeter.

Additionally, the pGL4.24 plasmid was restriction digested using the HindIII and KpnI restriction enzymes. This plasmid serves as the backbone and contains a minimal RNA polymerase II promoter driving expression of the firefly luciferase gene. The correct size of the digested plasmid (4300bp) was then verified by running a 1% agarose gel with a 100bp ladder (Fig.2.) This band was then cut from the gel for DNA extraction and purification. Finally, Gibson Assembly was used to ligate each enhancer to the restriction digested and purified pGL4.24 plasmid at the HindIII and KpnI sites. After selection of ampicillin-resistant colonies, plasmid DNA was extracted and sequenced to verify the correct sequence and orientation of insertion.

Once these plasmids were sequence-verified, they were then re-transformed into *sbl3* competent *E.coli* cells in order to generate more copies of the plasmids containing one of the three enhancer sequences. Finally, a midi-prep process was used to extract pure, supercoiled and endotoxin-free DNA from bacterial cells. These plasmids are used in further experiments with mammalian epithelial cells (MCF10a cells) and human colorectal cancer cells (HCT116 cells) expressing the two isoforms of p63 to verify that they perform enhancer functions.

A set of Luciferase assays were performed on two cell types: MCF10a (derived from normal human breast epithelial cells) and HCT116 (derived from human colon cancer cells). MCF10A cells natively express p63 while HCT116 cells do not. These experiments aimed to test the ability of p63 to activate enhancers and whether p63 would be sufficient for this activity. In order to do this, we created two cell models. First, we created an MCF10A cell line with reduced expression of p63, called MCF10a p63KD. A short hairpin RNA (shRNA) was designed against the p63 DNA binding domain, which is predicted to target all p63 mRNA transcripts. This was restriction cloned into a lentiviral vector backbone downstream of an RNA polymerase III-promoter. This vector also contained a selectable marke for puromycin resistance. A control shRNA was produced that targets an *Arabadopsis*-specific mRNA, such that it would not actually affect RNA stability in human cell lines. Virus was produced as described in Materials and Methods, and infected cells were selected using puromycin.

We then created a version of the HCT116 cell line that expresses either the ΔNp63 or the TAp63 isoform under control of a doxycycline (dox)-inducible promoter. ΔNp63 and TAp63 were subcloned into a lentiviral backbone plasmid downstream of a dox-inducible CMV promoter. This backbone also contained a cassette for the expression of the reverse Tet-transactivator protein, which allows inducible expression of ΔNp63 or TAp63 in the presence of

tetracycline analogs like dox. This backbone also contained a puromycin resistance gene for selection of infected cells with puromycin. Virion were produced and cells infected as described in Materials and Methods, and uninfected cells were removed from the population by selection with puromycin.

We then confirmed that the infected cell lines produced the phenotypes we desired. First, we performed a western blot experiment testing for p63 and p53 expression in MCF10a cells demonstrated that both proteins were expressed in this cell type and served as a positive control for the p63 antibody (Fig. 3A). We confirmed the construction of the inducible ΔNp63 and TAp63 HCT116 cell lines through western blots an antibody against a common region present in both p63 isoforms (Figure 3B). In both cases, we did not observe any expression in the absence of doxycycline, confirming the inducibility of the gene. We next measured expression of p63 in either control MCF10A or p63shRNA-expressing MCF10A cells after exposures to either a control (DMSO) or two different p53-activating drugs (Nutlin-3A or etoposide). The rationale for exposing the cells to the different drugs was demonstrate that p63 and p53 are regulated in different manners. Expression of p53 increases in response to these drugs (data not shown). As expected, cells expressing p63shRNA show reduced protein expression of p63 (Fig. 3C).

We then moved forward to test the effect of either removal of p63 (from MCF10A) or forced expression of p63 (in HCT116) on activity of our different enhancers. Enhancers 220, 231, and 245 were transfected (along with a control vector consitutively expressing Renilla luciferase) into MCF10a WT, MCF10a p63KD, and HCT116 WT cell lines, and were expressed in the HCT116 ΔNp63 cell line using 1 µg/µL Doxycycline. A reporter construct lacking an enhancer was used as a negative control in all Luciferase assays. The Luciferase results were first normalized to the expression of Renilla, which represents a control to measure transfection

efficiency across experimental conditions. All of the enhancers were then compared to the control “no enhancer” value (experimental/control). This allows for the comparison of enhancer activity and transfection efficiency across conditions and cell types.

The three enhancers we tested appear to have enhancer activity in MCF10A WT cells (where p63 is expressed natively), as their activity levels are higher than that of the no enhancer control (Fig. 4). Luciferase assays performed on the same enhancers in a MCF10a p63KD cell line shows a substantial decrease in their activity levels compared to MCF10a WT cells (Fig. 5). These differences between control and p63shRNA-expressing cells are statistically significant. This suggests that p63 is required for full activity of these enhancer elements. To determine if p63 expression is sufficient to drive this activity, Luciferase assays were performed on HCT116 WT cells (which do not natively express p63). The results show a reduction of enhancer activity compared to MCFT10A WT cells (Fig. 6). Furthermore, two of the three enhancers showed such low activity levels that they have the same or worse activity than the no enhancer control. This suggests that all of the enhancers are more active in MCF10A cells and that two (220 and 231) have activity in HCT116 cells. To determine if reduction in activity was due to the absence of p63 in the native cell environment, Luciferase assays were performed on HCT116 inducible-ΔNp63 cells, where expression of the ΔNp63 isoform is forced through the use of a doxycycline-responsive transcriptional unit. The results show that overexpressing p63 in HCT116 cells was not sufficient to activate any of the enhancers beyond the level we see in cells lacking p63 (Fig. 7). These data suggest that p63 expression alone is insufficient to drive the activity, but that it is required for enhancer function in MCF10A cells.

We next wanted to better understand what genes are regulated by p63-bound enhancers. Enhancers can be great distances from the regulated gene, therefore it is often difficult to

attribute an enhancers activity to a given gene. To further investigate the activity of these cluster 2 enhancers and the genes that they regulate, we selected an additional ten regions from the original MPRA experiment that were shown to be enhancers. We designed an experiment using utilizing CRISPR-dCas9-KRAB to block the activity of a given enhancer element and then measure the expression of thousands of genes in parallel using RNA-seq. CRISPR-dCas9-KRAB binds to any region of the genome that we desire and blocks critical transcription factors from interacting with the enhancer DNA sequences through formation of H3K9me3-enriched heterochromatin. The guide RNA sequences for dCas9 were designed to target dCas9 KRAB to the p63 binding sites of the 10 enhancers. We included both a positive control for a known enhancer at the GDF15 gene and a negative control, which was targeting an enhancer that is inactive in MCF10A cells (FGF2). Therefore, we could benchmark our results and determine whether the experiment was performed properly since we had an expectation of the results.

A lentiviral contruct was used to transduce MCF10A cells with a specific gRNA sequence and dCas9 KRAB, and infected cells were selected using puromycin. Upon establishment of stably-expressing cell lines, total RNA was isolated and used for polyA+ RNA selection using polydT-coupled magnetic beads. Barcoded RNA sequencing libraries were generated from the polyA+RNA for all 12 cell lines. Successful sequencing library generation was first verified through capilary electrophoresis (Fig.8), in which the presence of DNA between 200 and 600bp indicate the correct sizes of the RNA sequencing library. These libraries were then absolutely quantified using known quantities of DNA library using a qPCR-based approach and pooled together. The library was then sequenced in a 75 cycle/read reaction on an Illumina NextSeq500 at the Center for Functional Genomics. RNA-expression values

(Transcripts Per Million, TPM) were generated by aligning the sequenced RNA reads against the human hg19 RefSeq reference transcriptome using *salmon.*

The expression level of GDF15, our positive control, was dramatically reduced (≈20 fold) compared to FGF2 targeting (black, control) when the dCas9-KRAB complex was targeted to it's known enhancer positive element (Figure 9A). As evidence of specificity, the nearest 7 genes upstream and 7 genes downstream were unaffected by placement of dCas9-KRAB at the GDF15 enhancer (Figure 7A). We also saw no effect on nearby gene expression when targeting the FGF2 enhancer, globally (black dots, Figure 9A). When the dCas9-KRAB complex was then targeted to ten different cluster 2 enhancers, we only found clear evidence of enhancer activity when targeting dCas9-KRAB to enhancer 197. This led to the downregulation of three nearby genes, LINC01698, MIR205, and CAMK1G. We confirmed MIR205 expression was reduced when enhancer 197 was targeted by dCas9-KRAB using RT-qPCR compared to control targeting of dCas9-KRAB to the inactive FGF2 enhancer. This result is statistically significant (Figure 10). Overall, these data suggest that dCas9-KRAB-targeting to p63 bound enhancers coupled to RNA-seq can identify gene targets of those enhancers but that not all p63-bound enhancers might act as enhancers in all contexts.

Conclusions

A preliminary massively parallel reporter assay (MPRA) revealed a group of enhancers which only bound to p63 (Fig. 11) and were unable to be bound by the paralog p53. We aimed to identify whether p63 activity was required for enhancer function and whether p63 binding or expression alone would be sufficient to activate these enhancers outside of the native cell context. We then asked whether we could use a targeted enhancer inhibition approach (dCas9-KRAB) to identify the gene targets of p63-bound enhancers. We began by asking about the role of p63 at enhancers by studying three random p63-bound enhancers we chose from our MPRA screen. We clones the three enhancers from genomic DNA and successfully cloned them into a Luciferase reporter plasmid. We then used these enhancers to investigate whether p63 can bind and activate these enhancer regions. These plasmids were then transfected into HCT116 cells, which natively do not express p63, and demonstrated these enhancers had reduced or no activity. Expression of both isoforms of p63 in HCT116 cells suggests that p63 expression alone is insufficient to drive expression of these enhancers.

Our results focus on three different p63-bound enhancers, which our MPRA suggests might be p63-dependent. In MCF10a WT cells, these DNA sequences seem to have enhancer activity as suggested by previous assays. In order to determine whether p63 was required, the MCF10a p63KD cell line was created to modulate levels of p63. Luciferase assays performed on this cell line shows that the activity of the same three enhancers decreases significantly compared to the wild-type cell line. These results suggest that p63 is required for the activity of these enhancers.

To determine whether p63 alone was sufficient to drive the activity of these enhancers, we performed Luciferase assays on HCT116 WT cells. Compared to MCF10a WT cells, the

three cluster 2 enhancers all show decreased activity levels in HCT116 WT cells, suggesting some transcription factors required for MCF10A-level activity are missing. Alternatively, a repressor protein/proteins might be present in HCT116 that leads to the observed reduction in activity. Interestingly, two of the enhancers exhibited such low activity levels that they were equivalent or less than the no enhancer controls, suggesting that enhancer activity is strongly context and cell type-dependent. Furthermore, when we forced expression of ΔNp63 in HCT116 cells we do not observe an increase enhancer activity. This suggests that p63 alone is not sufficient to drive the activity of these cluster 2 enhancers.

These data allow us to put forth a potential model for the activity of p63 at enhancers. p63 involving the differing cellular environments and proteins expressed between the two cell types. MCF10a cells natively express p63 and are an epithelial cell type, while HCT116 cells do not express p63 and are a tumor-derived cell lines that is mesenchymal in nature. The most straightforward model is that p63 activity requires additional factors that are present in MCF10A cells but absent from HCT116. That is to say, the set of factors and cofactors expressed in MCF10A allows full enhancer activity. HCT116 colorectal cancer cells are derived from colon. It is possible that certain transcription factors or cofactors that are required to interact with the p63 network may be compromised compared to that of MCF10a cells, which are derived from non-cancerous epithelial cells. Additionally, there may be key epithelial factors present in MCF10a cells that drive certain enhancers to work better. Further work can be done to examine the exact cofactors that work with p53- and p63-dependent enhancers in MCF10as, and to compare the activity of these enhancers in both cell types. One option is to perform saturating mutagenesis approaches to identify additional DNA sequences within the enhancer regions that might bind other required transcription factors. Additionally, we could perform DNA

footprinting experiments to infer the occupancy of different DNA elements within the enhancer. Finally, and perhaps most importantly, a comprehensive yeast-1-hybrid experiment can be performed against a bait of the p63-bound enhancer against a prey library of human transcription factors. This would allow the unambigous identification of the factors binding. This approach is limited by the difficulty of obtaining a genome-scale transcription factor library and by the non-native context of the assay (yeast versus intact human cell lines).

We then also sought to investigate the target genes of these p63-dependent enhancers. This question is perhaps the most difficult in all of transcriptional regulation, due to the highly contextual nature of enhancers. What we mean by this is that one gene can have many enhancers, and these enhancers can be redundant or context-dependent. Therefore, if we delete or block a redundant enhancer, we would be unable to identify the gene target. Further, because enhancers can be located incredily far from their target gene, proximity-based approaches (such as only looking at the nearest gene) tend to fail more often than not. We decided to answer this question *in vivo* using CRISPR-KRAB dCas9, which rapidly blocks enhancer activity. If we block the enhancers activity, then we presumably affect the linked gene. In order to measure the gene without prior knowledge of which gene could be affected by the enhancer, we used a transcriptome-wide measurement technique called RNAseq. This allowed us to target dCas9-KRAB to any enhancer and then measure the effect on the expression of all genes in the genome in an unbiased fashion.

Ten putative enhancers were chosen at random from our list of regions that were active in MCF10A cells in the MPRA assay. All of these enhancers showed high p63 occupancy. The KRAB domain facilitates the assembly of the proteins that generate heterochromatin, which blocks the activity of the enhancer by preventing the recruitment and binding of critical

transcription factors and co-factors. Using this approach, we discovered that enhancer 197 likely controls expression of up to three different genes nearby. Excitingly, one of these genes is MIR205, a microRNA that appears to be important for epithelial integrity. This is interesting because regulation of epithelial integrity is the same phenotype shared with p63. Our work provides a direct link between a master regulator of epithelial biology (p63) and a microRNA that is required for the same process. There are additional genes regulated by enhancer 197, including a long noncoding RNA (lncRNA) with as of yet undetermined function. Enhancers often work in hubs, where one enhancer might regulate expression of multiple promoters. Therefore, future work could investigate whether this enhancer regulates multiple genes directly or whether the reduction in expression of three genes is an indirect effect of the reduced expression of one of those genes. Overall, our work provides evidence for the enhancer function of this p63-dependent enhancer. Surprisingly, we did not see evidence for gene regulatory activity of the other p63-dependent enhancers. This can be for a number of biological and technical reasons. First, p63 has been shown to be repressive, therefore repressing this region with heterochromatin or with p63 binding might lead to the same functional outcome. RNAseq is very sensitive, but we did not perform multiple independent biological experiments, which might help us resolve whether small differences are in fact legitimate. Finally, we know enhancers oftentimes are redundant and work with other enhancers. In this case, blocking a redundant enhancer would not provide evidence of regulation unless the other enhancer was blocked.

We propose a model in which p63 works *in vivo* along with other cofactors in a complex regulatory network to drive transcription of downstream target genes responsible for epithelial morphogenesis. Our findings have enabled us to begin to decipher the complex gene regulatory networks that the p53 family of transcription factors is involved in. By examining the activity of

p63-bound enhancers in different native cell environments, we begin to understand what defines an enhancer, what cofactors an enhancer requires, and what transcription factors such as p63 do when they bind to their enhancer. Our results have also paved the way for distinguishing the characteristics of p63-bound enhancers from p53 enhancers, which appear to be more active and functional only after p53 stabilization. Further experiments can then be performed on this network to investigate other transcription factors that act in combination with these enhancers, as well as the other target genes of p63-dependent enhancers in addition to MIR205. Ultimately, understanding how the p63 network regulates gene transcription leads to a better understanding of the dynamics of these complex regulatory regions.

References

Biscontini, T. 2019. *Gene*. Salem Press Encyclopedia of Science.

Brady, C., Attardi, L., 2010. p53 at a glance. *Journal of Cell Science*, 123, 2527-2532. doi: 10.1242/jcs.064501

Koster, M., 2010. p63 in skin development and ectodermal dysplasias. *The Journal of Investigative Dermatology*, 30(10), 2352-2358. doi: 10.1038/jid.2010.119

Nostrand, J., Bowen, M., Vogel, H., Barna, M., Attardi, L., 2017. The p53 family members have distinct roles during mammalian embryonic development. *Cell Death & Differentiation*, 24, 575–579. doi: 10.1038/cdd.2016.128

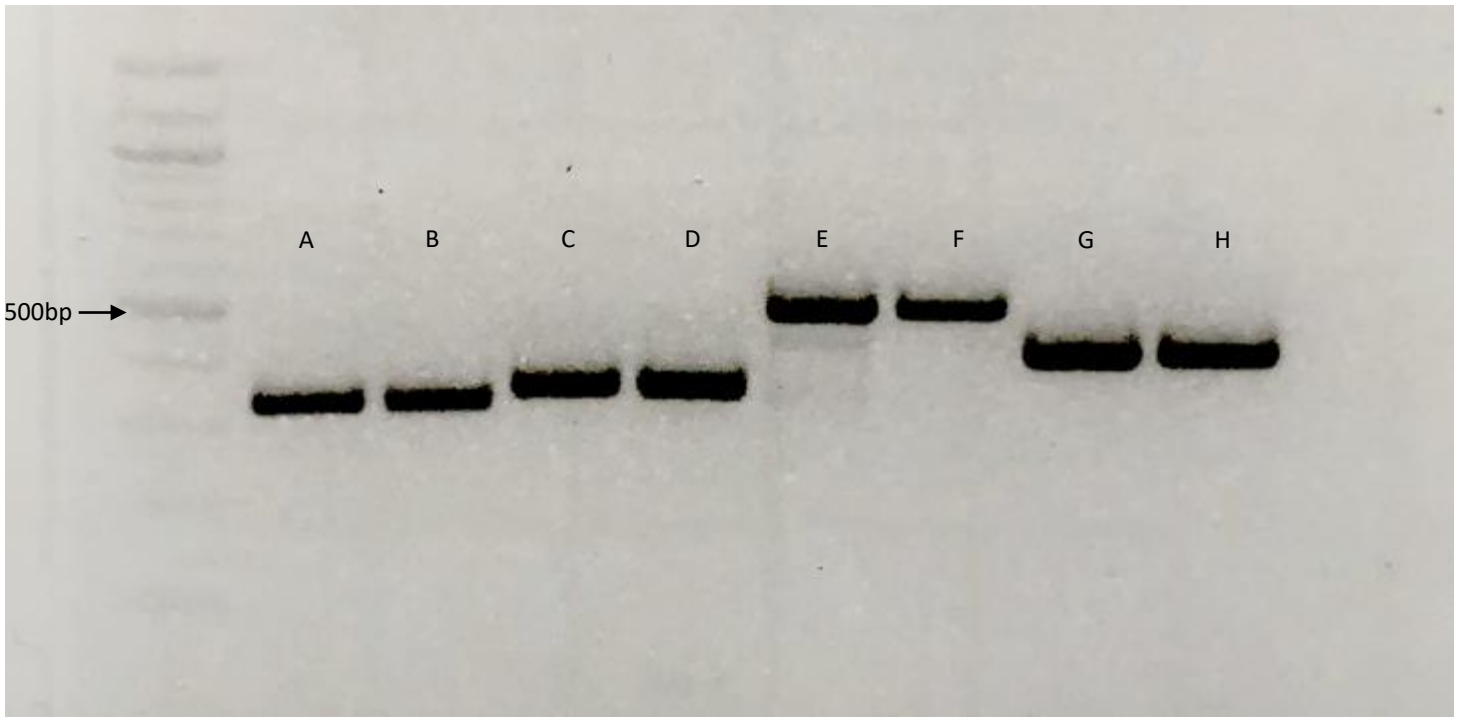
Romano, R., Smalley, K., Magraw, C., Serna, V., Kurita, T., Raghavan, S., Sinha, S., 2012. ΔNp63 knockout mice reveal its indispensable role as a master regulator of epithelial development and differentiation. *Development*, 139(4), 772–782. doi: 10.1242/dev.071191

Spitz, F., Furlong, E., 2012. Transcription factors: from enhancer binding to developmental control. *Nature Reviews*, 13, 613-626. doi:10.1038/nrg3207

Yeo, NC, Chavez, A, Lance-Byrne, A, Chan, Y, Menn, D, Milanova, D, Kuo, CC, Guo, X, Sharma, S., Tung, A, Cecchi, RJ, Tuttle ,M, Pradhan, S, Lim, ET, Davidsohn, N, Ebrahimkhani, MR, Collins, JJ, Lewis, NE, Kiani, S, Church, GM. Nat Methods. An enhanced CRISPR repressor for targeted mammalian gene regulation. 2018 Jul 16. doi: 10.1038/s41592 018-0048-5. 10.1038/s41592-018-0048-5

Zilfou, J., Lowe, S., 2009. Tumor Suppressive Functions of p53. *Cold Spring Harbor Perspectives in Biology*, 1(5), a001883. doi: 10.1101/cshperspect.a001883

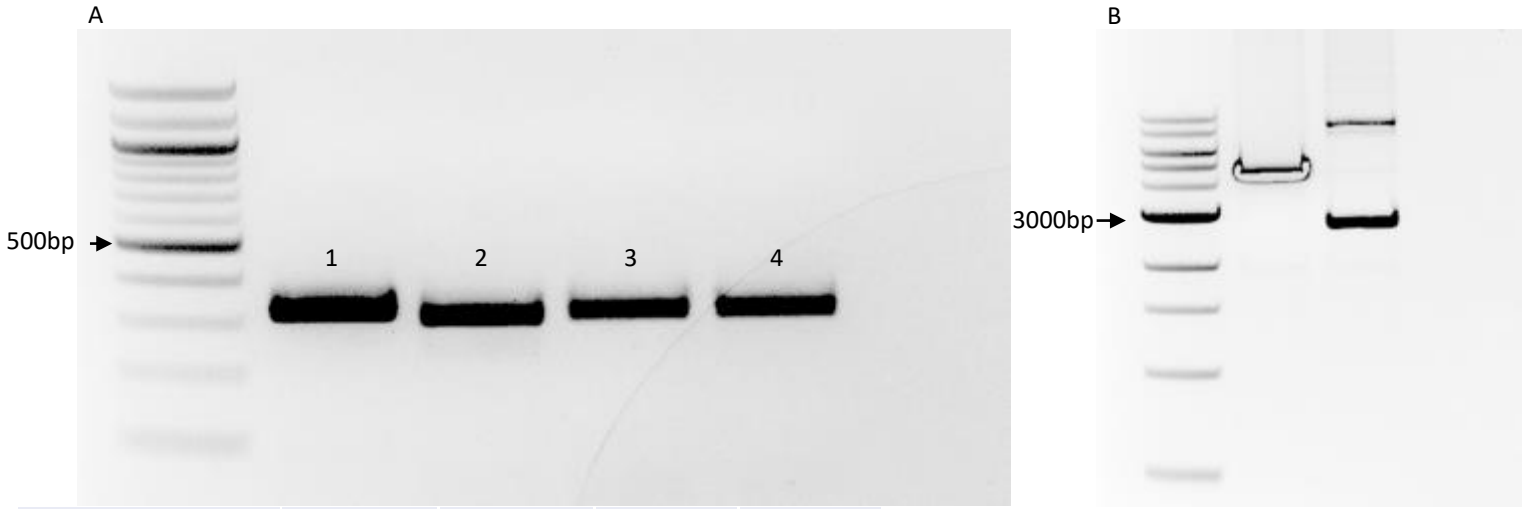
Figures



	A	B	C	D	E	F	G	H
template DNA (µL)	2	3	2	3	2	3	2	3
primers (µL)	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25
mastermix (µL)	12.5	12.5	12.5	12.5	12.5	12.5	12.5	12.5
water (µL)	9.25	8.25	9.25	8.25	9.25	8.25	9.25	8.25

Figure 1: Amplifying p63 enhancers out of the genome

PCR reaction run on a 2% agarose gel with a 100bp ladder. Lanes A and B contain reactions to amplify E186, lanes C and D to amplify E245, lanes E and F to amplify E231, and lanes G and H to amplify E220. Each pair of lanes contained either 100 or 150 ng of genomic DNA to optimize enhancer amplification.



	1	2	3	4
template DNA (µL)	1	1	1	1
primers (µL)	1.25	1.25	1.25	1.25
mastermix (µL)	12.5	12.5	12.5	12.5
water (µL)	10.25	10.25	10.25	10.25

	(+) control	(-) control
DNA (µL)	3.88	3.88
KPNI (µL)	1	0
HindIII (µL)	1	0
buffer (µL)	5	5
water (µL)	39.12	41.12

Figure 2: Performing Gibson assembly to ligate p63 enhancers with pGL4.24

(A) Gibson primers were first ligated to the four enhancers and then run on a 2% agarose gel with a 100bp ladder. Lane 1 corresponds to E186, lane 2 to E245, lane 3 to E231, and lane 4 to E220. (B) pGL4.24 was restriction digested at the HINDIII and KPNI sites, and the cut plasmid was run on a 1% agarose gel with a 1kb ladder.

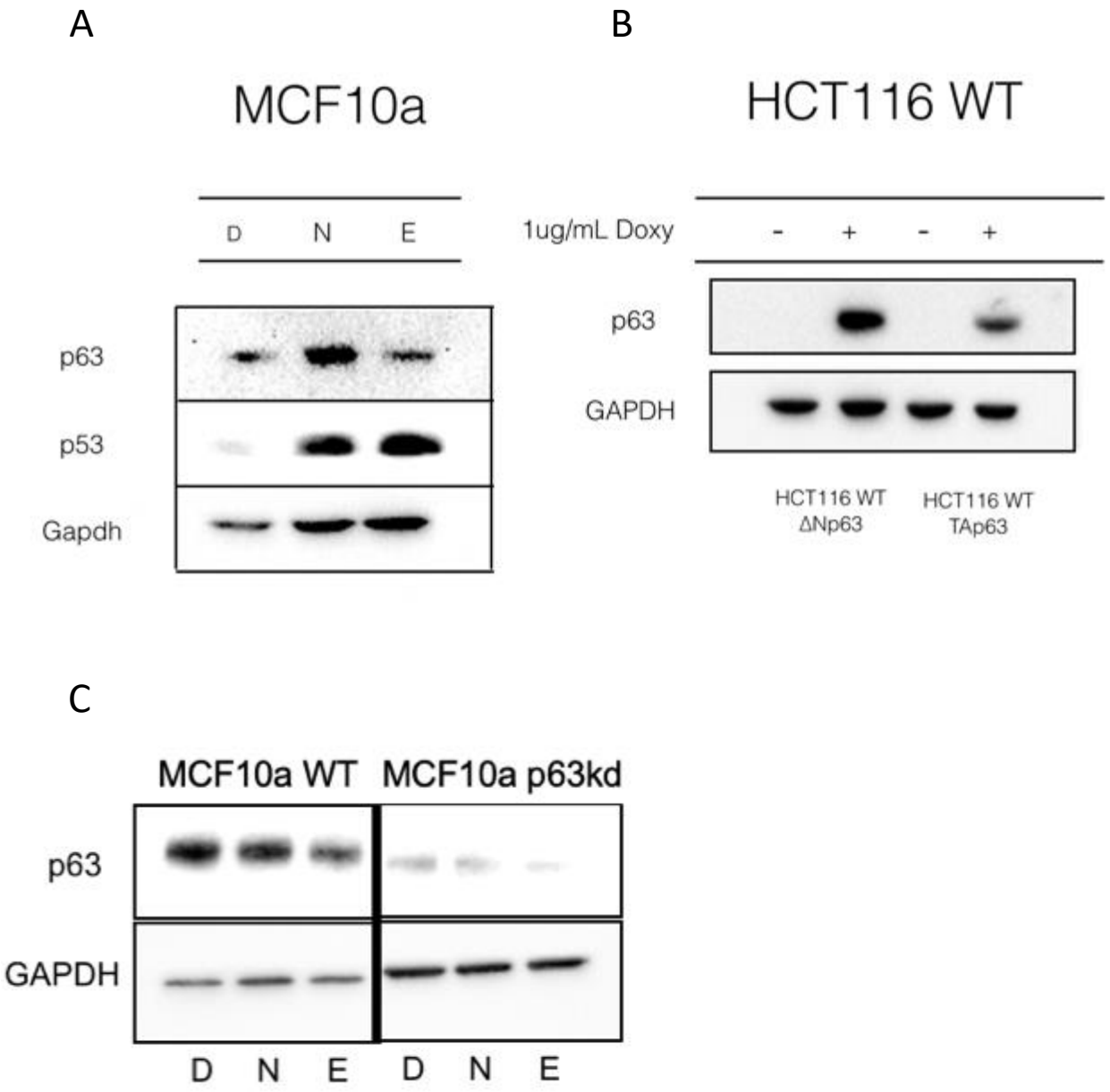


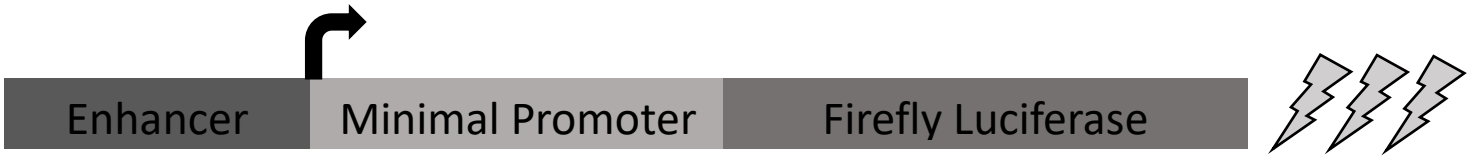
Figure 3: p63 is expressed in MCF10a and HCT116 WT cells, and is not expressed in MCF10a p63KD cells

(A) p63 is expressed in MCF10a cells in the presence of DMSO, 5 μM Nutlin and 100 μM Etoposide. p53 is expressed in presence of 5μM Nutlin and 100μM Etoposide.

(B) ΔNp63 and TAp63 were expressed in HCT116 WT cells in the presence of 1 μg/mL Doxycycline.

(C) p63 expression is greatly decreased in MCF10a p63KD cells. This was achieved by targeting shRNAs to p63 mRNA transcripts to target them for destruction by host machinery. GAPDH was used as a control in all three experiments

A



B

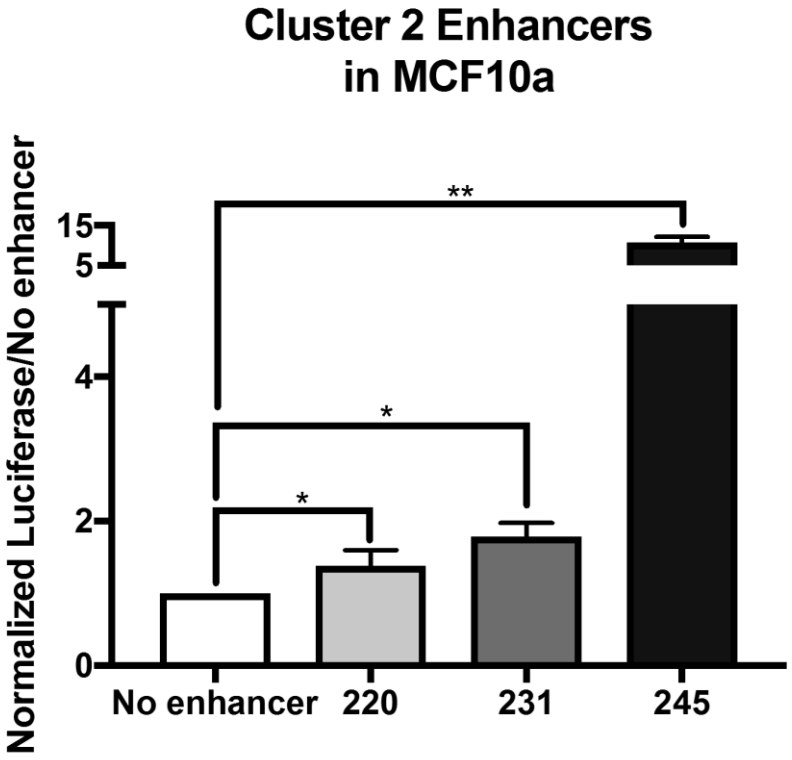


Figure 4: Cluster 2 enhancers are active in MCF10a cells
(A) Luciferase assay schematic.
(B) Enhancers 220, 231, 245 expressed in MCF10a cells.

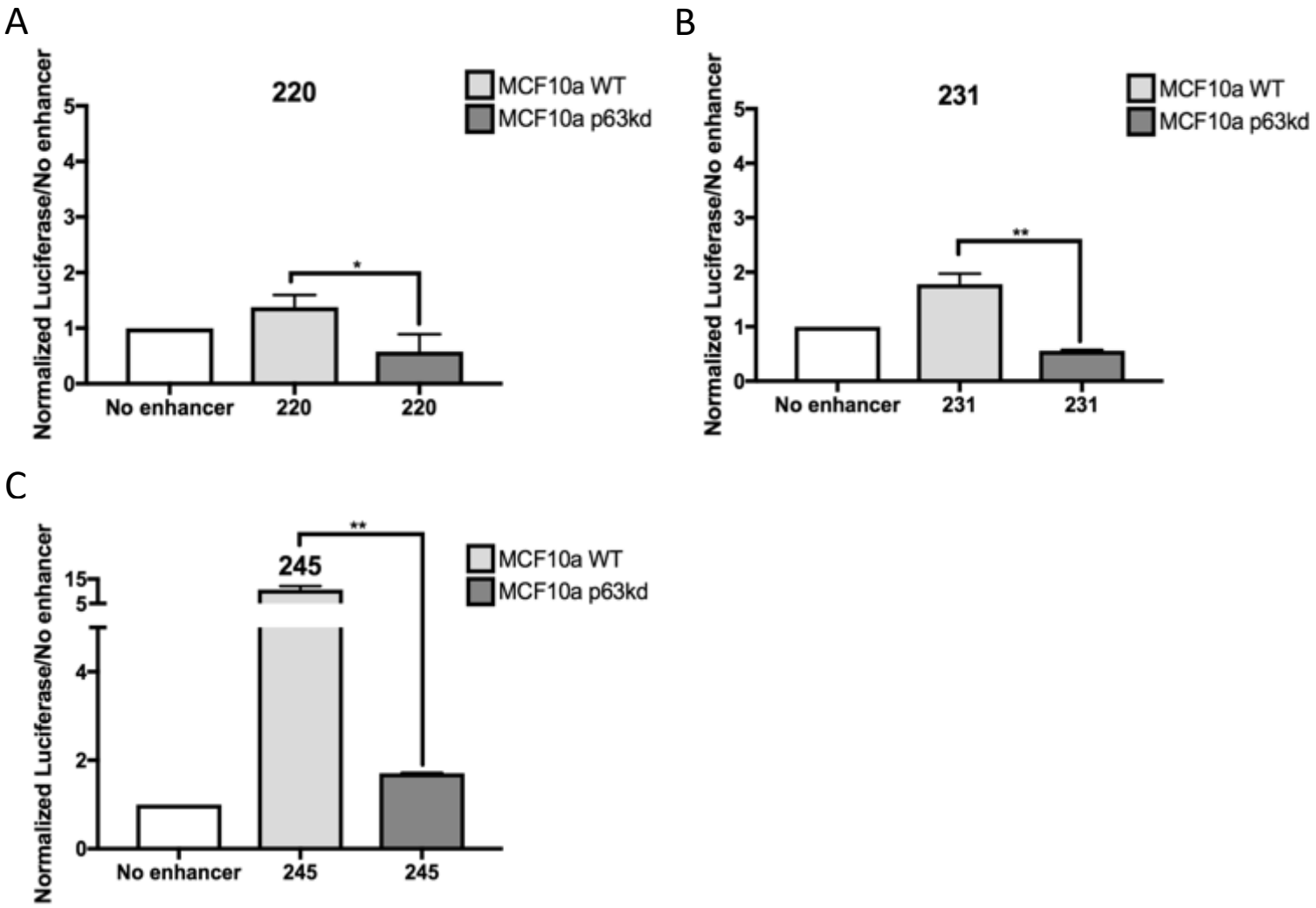


Figure 5: Reduced cluster 2 enhancer activity in MCF10a p63KD cells compared to MCF10a WT cells

(A) Enhancer 220 expressed in MCF10a p63KD cells vs. MCF10a cells.
(B) Enhancer 231 expressed in MCF10a p63KD cells vs. MCF10a cells.
(C) Enhancer 245 expressed in MCF10a p63KD cells vs. MCF10a cells.

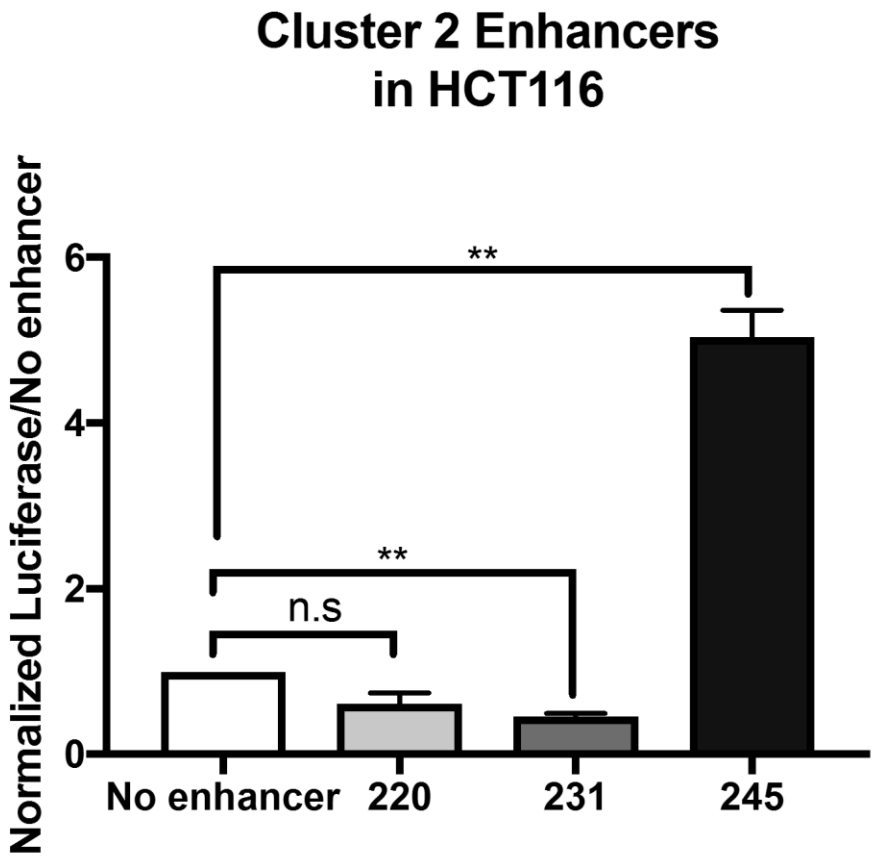


Figure 6: Cluster 2 enhancers show low activity in HCT116 WT cells
Enhancers 220, 231, 245 expressed in HCT116 WT cells.

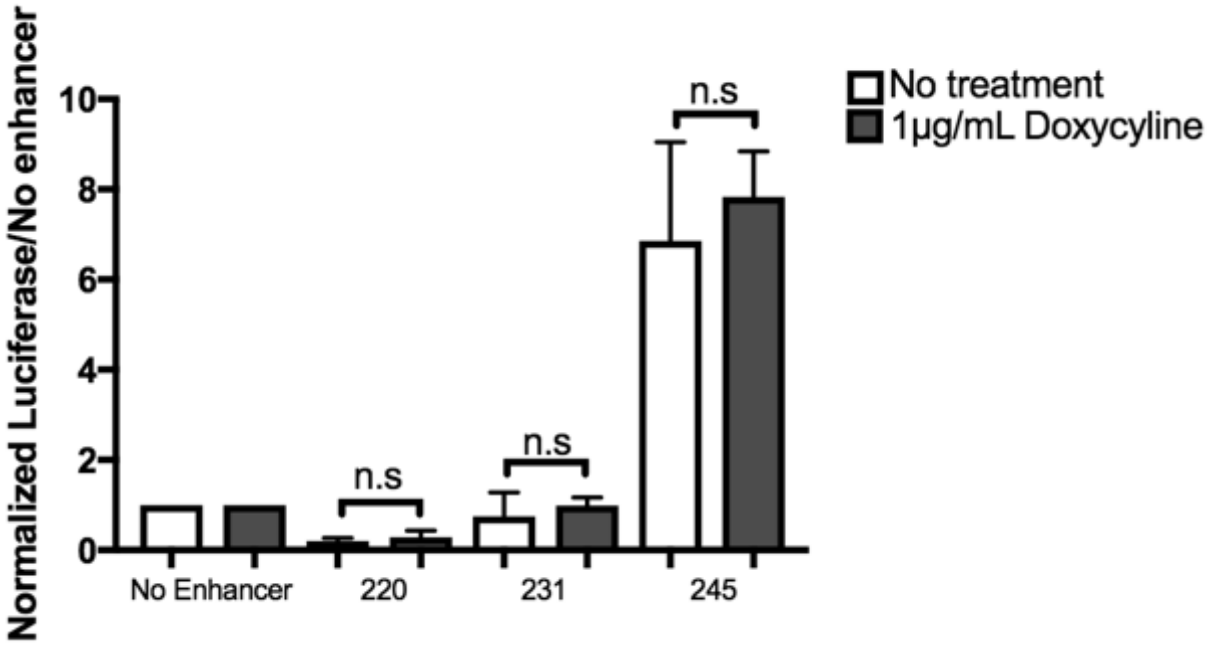


Figure 7: Expressing p63 in a non-native cell context is not sufficient to increase enhancer activity

Enhancers 220, 231, 245 expressed in HCT116 ΔNp63 cells vs. HCT116 WT cells. ΔNp63 was expressed in HCT116 WT cells in the presence of 1 µg/mL Doxycycline.

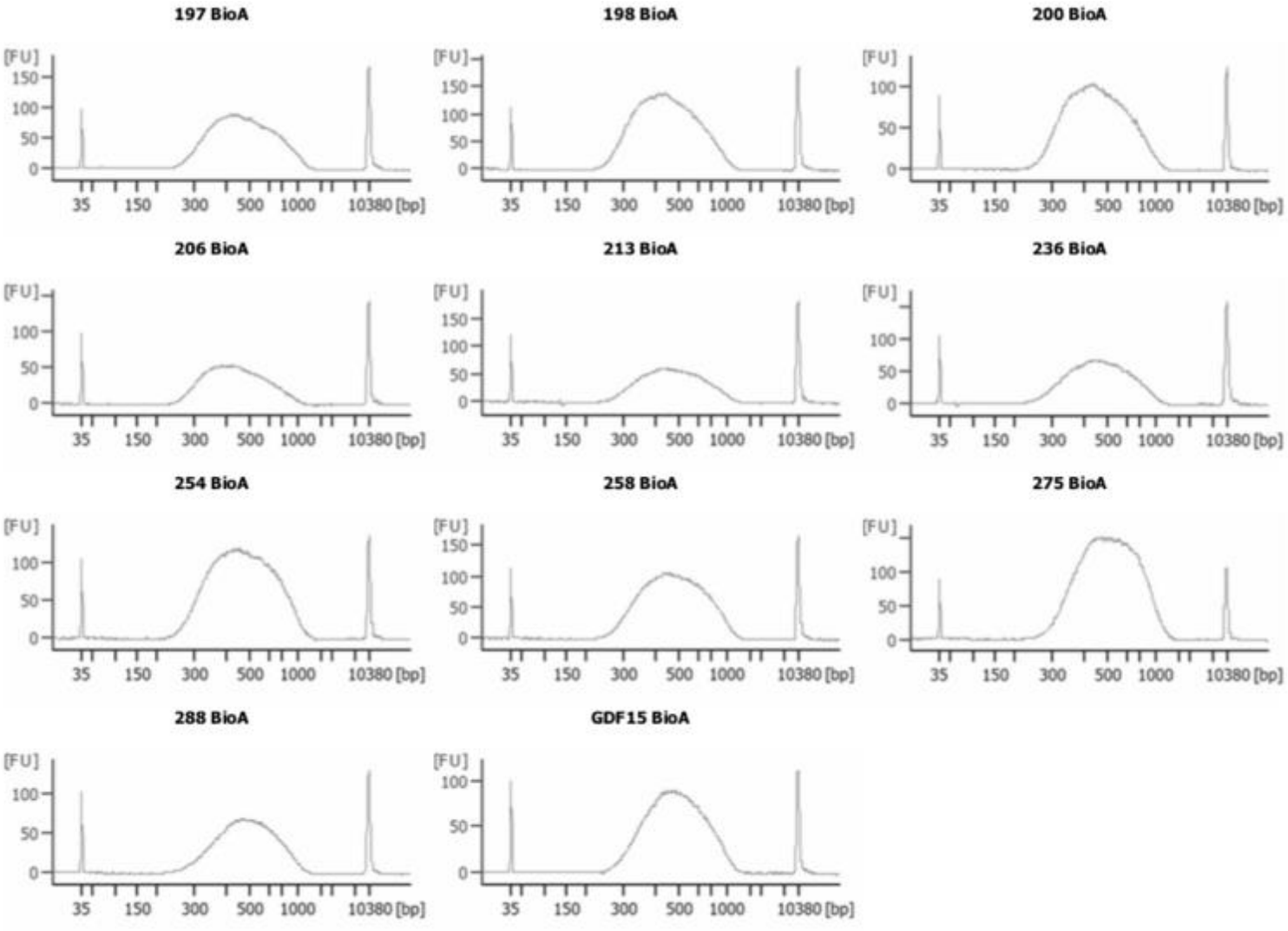


Figure 8: CRISPR-KRAB dCas9 libraries are of the appropriate size and composition for sequences

The presence of peaks verifies the correct sizes of the RNA sequencing library. Ten enhancers were studied along with a positive control, GDF15 and FGF2.

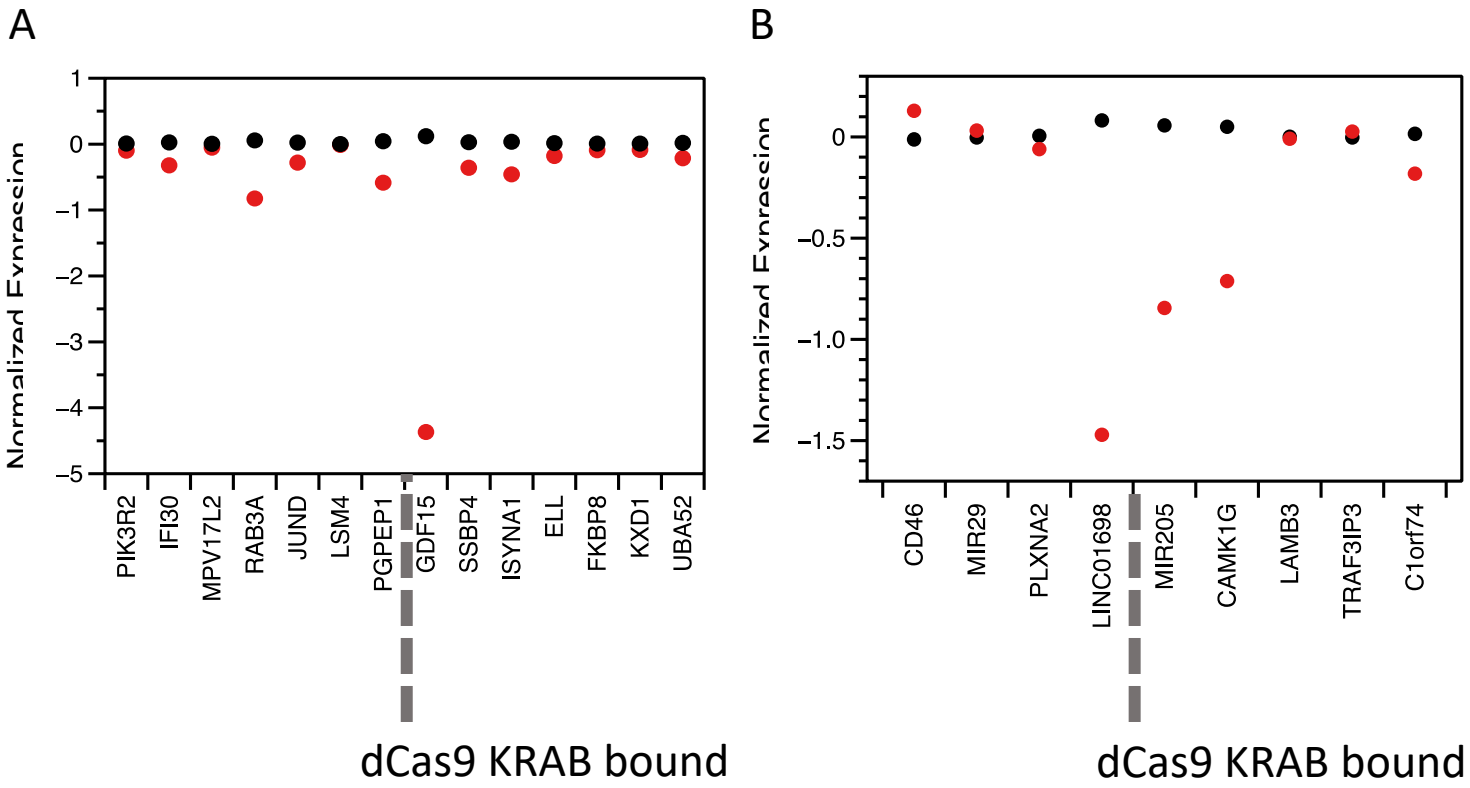


Figure 9: RNA-seq of KRAB-bound enhancer 197 reveals the downregulation of 3 downstream genes

(A) The GDF15 enhancer was targeted by dCas9 KRAB as a control.

(B) Targeting dCas9 to enhancer 197 causes the expression of LINC01698, MIR205, and CAMK1G to decrease.

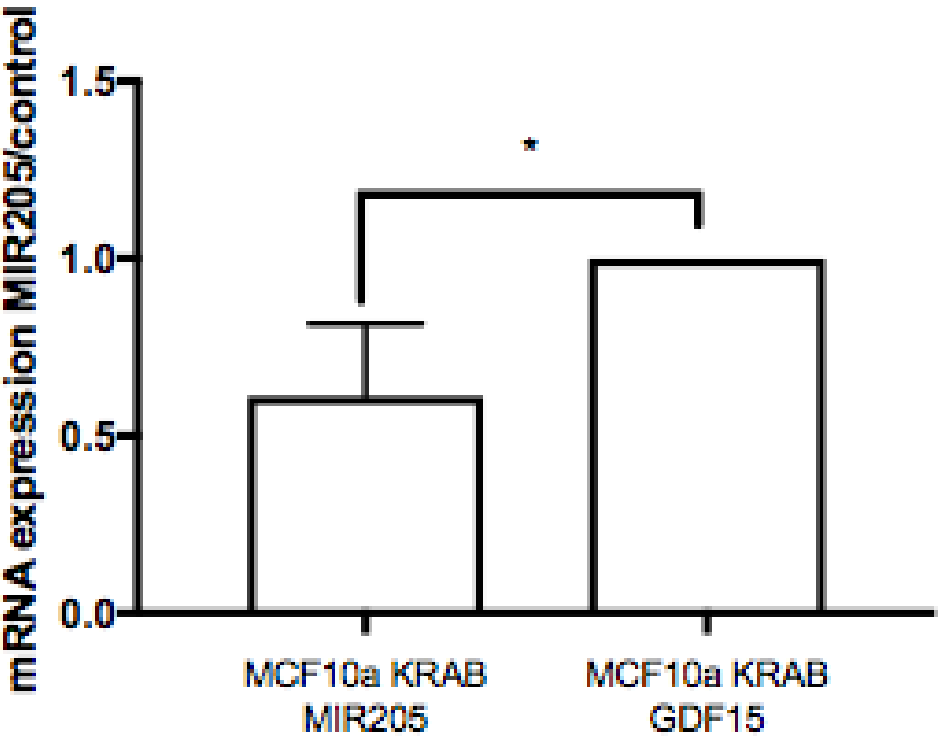


Figure 10: qPCR of MIR205 gene

qPCR was performed on the MIR205 gene after dCas9-KRAB was targeted to enhancer 197.

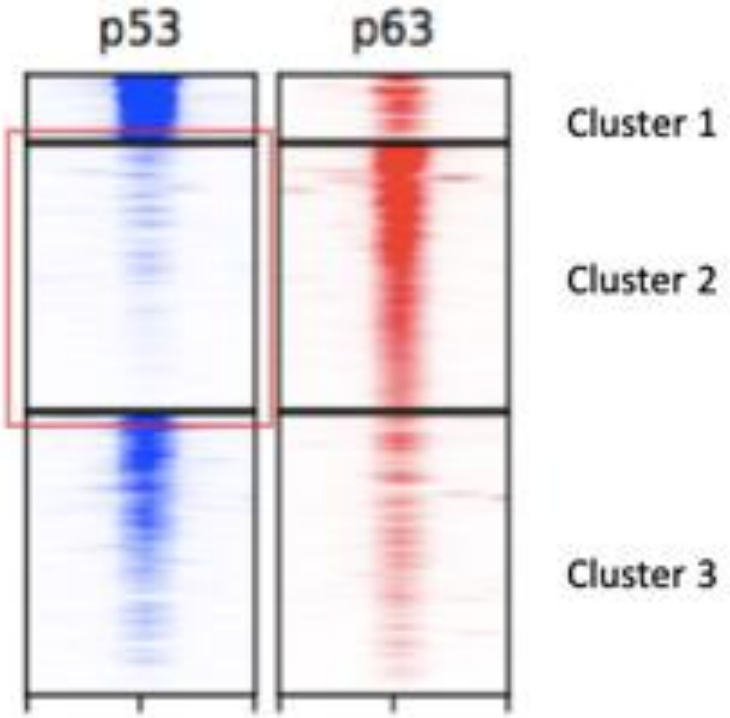


Figure 11: Enrichment of p53 and p63 present at all p63 enhancers sequences screened in MPRA

A preliminary MPRA screen of 296 enhancers. The p63-dependent cluster 2 enhancers of interest are boxed in red.