UNIVERSITY OF CALIFORNIA
RIVERSIDE

# An Investigation into Pumilio's Interactions with the miRNA Machinery and its Role in Cell Adhesion and Migration

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

Doctor of Philosophy

in

Genetics, Genomics, and Bioinformatics

by

Erin L. Sternburg

March 2019

Dissertation Committee:
Dr. Fedor Karginov Chairperson
Dr. Xuemei Chen
Dr. Frances Sladek

The Dissertation of Erin L. Sternburg is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

It is important to note that this work happened with the help of many colleagues, professors, family members, and friends. I would like to express my deepest appreciation and gratitude to my PhD advisor, Dr. Ted Karginov, whose intellectual rigor is matched only by his good nature. I am truly fortunate to have had the opportunity to work with him.

Many thanks to the past and present members of the lab, Jason Estep, Yahui Li, Kris Dias, Daniel Nguyen, and Jordan Lillibridge, for being the best lab family I could have asked for. It has been a pleasure learning from and with you, and I will always value your friendship.

I thank the members of my dissertation committee, Dr. Xuemei Chen and Dr. Frances Sladek, whose insights and guidance through the years were instrumental in the development of these projects. I also thank the following UCR faculty for their guidance in the first two years of my PhD training: Dr. Weifeng Gu, Dr. Thomas Girke, Dr. Hailing Jin, Dr. Sika Zheng, Dr. Emma Wilson from my qualifying exam committee, and Dr. Paul Larsen from guidance committee.

Thanks to Martin Riccomagno and Dr. Marianne Bronner for antibodies, and to Dr. Carolyn Rasmussen for use of the confocal microscope, I would also like to thank Rattapol Phandtong and Dr. Prue Talbot for all their help and advice regarding image analysis.

Finally, I would like to thank my family. To my dad, whose scientific mind has been an inspiration to me throughout my life. Thank you for instilling in me a lifelong love of science. To my mom, whose compassion and kind heart have helped me through good days and bad. Thank you for always checking in on me and always listening. To my brother, Eric, for reminding me to relax and enjoy the little moments in life. And to Jason, for making me smile every day and for always being excited for life's adventures. Words cannot express how much I love you all.

In accordance with graduate program guidelines, Chapter 2, in full, is a reprint of a published paper in the Journal of Scientific Reports (DOI: 10.1038/s41598-018-33596-4). 

# AUTHOR CONTRIBUTIONS

## CHAPTER 1

Erin Sternburg and Fedor Karginov wrote and edited the manuscript.

## CHAPTER 2

Erin Sternburg performed experiments, analyzed data, wrote and edited the manuscript. Jason Estep, Daniel Nguyen and Yahui Li performed experiments. Fedor Karginov conceived the study, analyzed data, wrote and edited the manuscript.

## CHAPTER 3

Erin Sternburg conceived the study, performed experiments, analyzed data, wrote and edited manuscript. Fedor Karginov conceived the study, analyzed data, wrote and edited the manuscript.

## CHAPTER 4

Erin Sternburg wrote and Fedor Karginov edited the manuscript.

ABSTRACT OF THE DISSERTATION

An Investigation into Pumilio's Interactions with the miRNA Machinery and its Role in Cell Adhesion and Migration

by

Erin L. Sternburg

Doctor of Philosophy, Graduate Program in Genetics, Genomics, and Bioinformatics
University of California, Riverside, March 2019
Dr. Fedor Karginov, Chairperson

RNA-binding proteins (RBPs) profoundly impact mammalian cellular function by controlling distinct sets of transcripts, often using sequence-specific binding to 3′ untranslated regions (UTRs) to regulate mRNA stability and translation. In addition to individual effects of RBPs, there are also examples of co-regulatory interactions between multiple RBPs occupying the same 3'UTR. Two well-characterized and highly conserved RBPs, Argonaute2 (AGO2) and Pumilio (PUM1 and PUM2), are known to bind overlapping sets of transcripts and individual examples of cooperative interactions between the proteins have been described. To further assess the extent of co-regulation between these proteins, transcriptome-wide changes in AGO2-mRNA binding upon PUM knockdown were quantified by CLIP-sequencing. The presence of PUM binding on the same 3′ UTR corresponded with both cooperative and antagonistic effects on AGO2 occupancy. In addition, PUM binding sites that overlap with AGO2 showed differential, weakened binding profiles upon abrogation of AGO2 association, indicative of cooperative interactions. In luciferase reporter validation of candidate 3′ UTR sites where AGO2 and

PUM colocalized, three sites were identified to host antagonistic interactions, where PUM counteracts miRNA-guided repression. We also characterized a new role of both PUM1 and PUM2 in regulating cell adhesion and migration. PUM double knockout (DKO) T-REx-293 cells grew in clumps, which arose from an inability to escape cell-cell contacts. In addition, defects in collective cell migration and actin morphology were also observed. RNA-sequencing further validated defects in gene categories related to adhesion and migration. In total, this work highlights the importance of further investigation into the biological roles of RBPs as well as the complex interactions between them.

# TABLE OF CONTENTS

LIST OF FIGURES

**CHAPTER 1**



**CHAPTER 2**

**CHAPTER 3**



**APPENDIX C**

LIST OF TABLES

**CHAPTER 3**



**APPENDIX A**



**APPENDIX B**



**APPENDIX C**

LIST OF SUPPLEMENTAL VIDEOS

# CHAPTER 1

# Global Approaches to Studying RNA-Binding Protein Interaction Networks


## Abstract

RNA-binding proteins (RBPs) execute a wide range of functions that are important in almost all aspects of cellular biology. RBPs are often influenced by their local environment, which includes interactions with the RNA landscape as well as with other RBPs. In recent years, the use of global approaches as a means of studying RBP interactions has become possible through the development of genome-wide computational and experimental methods. These RBP interaction studies have provided evidence not only to the extent of these interactions, but to their complexity as well. In addition, there is growing evidence to support the existence of higher order RBP regulatory networks that drive cellular processes. This chapter will summarize the current literature on RBP interactions, address how we approach studying them, and form a global picture of these ribonucleoprotein networks.

## From RBP basics to RBP networks

RNA-binding proteins (RBPs) are a broad class of proteins that orchestrate most essential cellular processes. Some RBP-RNA interactions form stable ribonucleoprotein (RNP) particles with defined roles, while other RBPs (and RNPs) interact transiently to process RNA, regulate RNA function, and control RNA fate in the cell. Importantly, RBPs are the key players in post-transcriptional regulation of both messenger RNAs (mRNAs) and non-coding RNAs, encompassing RNA splicing, transport, modification, stability, and translation[1,2]. Many of these roles and structural/sequence features of RBPs are deeply conserved across eukaryotes [3]. RBPs directly interact with RNA through sequence-specific, structure-specific, and nonspecific binding modes [4-7]. In addition, target specificity for a subclass of RBPs is achieved by association with microRNAs[8-10].

RBPs play a large role in cellular processes that impact development and homeostasis. mRNA expression during development is under pervasive spatial and temporal regulation. In a large part, this precise regulation is accomplished by modulation of RBP activity, which can affect changes in gene expression faster than transcription[11]. Similarly, RBPs can accomplish fine-tuned regulation of mRNA targets in order to quickly respond to external or internal stimuli. Because of their widespread functions, misregulation of RBPs is often associated with disease[12-15].

To illuminate the molecular mechanisms that underpin the cellular roles of RBPs, researchers have been systematically addressing several key types of questions about the binding and impact of individual RBPs, and their interplay with the RNA landscape. Determining RBP effects on mRNA stability and splicing patterns is informative to their

molecular modes of action. These measurements have typically involved knockdown/knockout/overexpression followed by RNA-seq or microarray analysis[16]. Most RBPs function through transient recruitment of downstream pathway components or in stable larger complexes, and the nature of these interactions have been investigated by *in vitro* and *in vivo* protein-protein interaction studies[17-19]. Additionally, it is important to understand how the RNA landscape affects RBP function and binding preferences, and *vice versa*. To this end, methods that develop a transcriptome-wide picture of RNA modifications and secondary structure have aimed to tackle these questions[20-24] (Figure 1.1C).

In addition to the above, a central question in these studies is determining the binding specificity and the set of targets of a given RBP. Building on earlier methods, *in vitro* binding assays coupled to high-throughput sequencing or microarrays have uncovered the binding preferences for many RBPs [25,26] (Figure 1.1A). *In vivo* targeting has been initially probed by RBP immunoprecipitation followed by microarray or sequencing (RIP-chip or RIP-seq), thus identifying whole transcripts that are bound by RBPs [27-29]. Related methods that use crosslinking followed by immunoprecipitation (CLIP), limited RNase digestion, and sequencing, have been even more broadly applied (Figure 1.1B). These approaches focus in on a short fragment of associated RNA, allowing for the determination of precise binding sites [30-35]. *De novo* motif enrichment analysis of the associated fragments can then subsequently reveal the sequence-specific binding preferences[36]. . While a substantial number of RBPs have been systematically characterized in this manner[37], many remain unstudied to date.

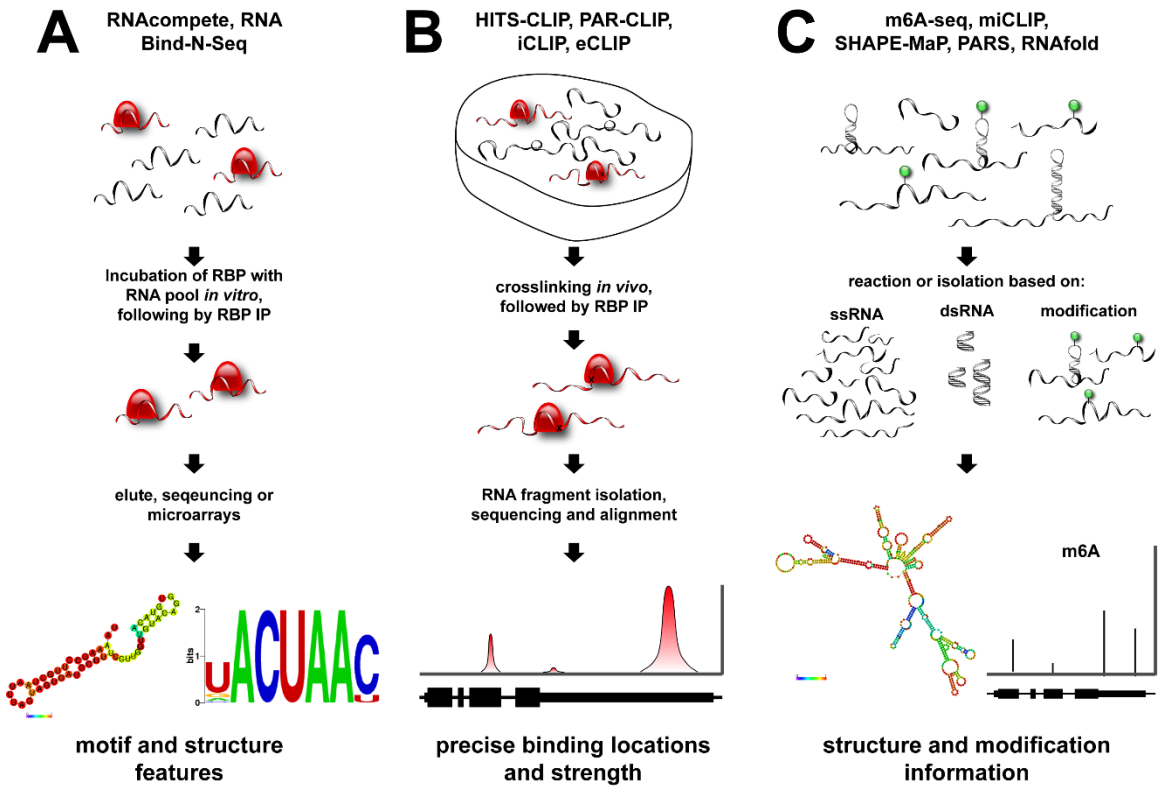



**Figure 1.1: Schematic of common RBP and RNA characterization methods.** A) In vitro RNA motif identification. B) In vivo RBP binding site identification via CLIP-sequencing. C) RNA structure and modification identification.

In addition to the static binding picture, it is essential to determine how RBP targeting changes throughout development, in response to an acute change in cellular environment, as well as in disease. For these, CLIP-seq is a key experimental resource to measure RBP binding in multiple tissues types or pathological states such as cancer. Understanding differential targeting can be instrumental in identifying the biologically relevant interactions under specific conditions, providing key connections to downstream phenotypes. Overall, measurement of RBP expression, cellular localization, changes in RNA landscape and binding partners, and the shifting RBP binding preferences, are fundamental to understanding their regulatory roles.

Finally, while the focus of most studies to date has been on the functional roles of individual RBPs, it is increasingly recognized that higher-order interactions between RBPs on specific mRNAs are common, and these interactions mediate pervasive and complex regulatory outcomes. Throughout their lifecycles, mRNAs are continuously associated with large and dynamically changing sets of RBPs that contribute to their regulation in a non-linear fashion. In addition, recent studies hint that RBPs form large regulatory networks which act to coordinate complex cellular processes. Thus, much of the leading-edge research is concentrated on determining how RBPs interact with one another to enact combinatorial regulation. These studies require new sets of increasingly integrative, transcriptome-wide approaches, and have begun to expose the principal concepts of this broad regulatory paradigm of post-transcriptional control, as described below.

## Interaction types and mechanisms gathered from studies on specific mRNAs

Examination of interactions between particular RBPs and/or miRNAs on specific mRNAs have delineated some of the mechanisms of such co-regulatory events and their effects on expression outcomes, as detailed in recent reviews [38-40]. The interactions can be cooperative, where the binding of one RBP leads to increased binding and/or activity of its partner RBP[41-58]. Alternatively, an RBP can destabilize the binding or repress activity of another, resulting in an antagonistic interaction[47,59-73].

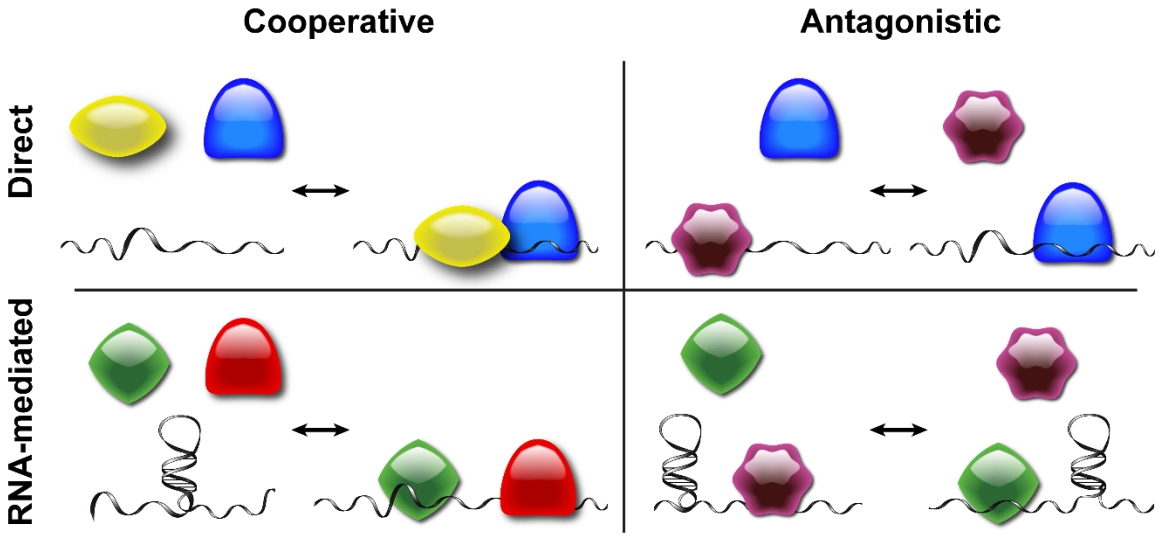



**Figure 1.2: Mechanisms of cooperative and antagonistic interactions between RBPs.** Direct association of two RBPs can lead to co-recruitment or steric hinderance. RBPs can also indirectly affect one another through cooperative or antagonistic RNA-mediated interactions.

Several mechanisms have been proposed to mediate these interactions. RBPs with overlapping binding sites may antagonize each other due to direct steric hindrance (Figure 1.2, top right panel). For example, expression of Dnd1 is thought to counteract miRNA-mediated repression by binding to sites that overlap with the miRNA seed, making it unavailable[59]. RBPs that occupy nearby mRNA sites have also been observed in cooperative interactions, where their co-recruitment is likely enhanced by direct binding between the RBPs (Figure 1.2, top left panel). Such direct binding has been demonstrated in some cases by co-IP or MS approaches, but is expected to be fairly weak and context-dependent, since those RBPs are not in stable stoichiometric complexes[47,49]. Alternatively, interactions can be enabled by changes in RNA secondary structure, where binding of one RBP can either mask or open the binding sites for others (Figure 1.2, bottom panels). Through this mechanism, binding of Pumilio to a target mRNA unfolds nearby secondary

structures, which allows for the subsequent binding of miRNAs[41,74-76]. The latter mechanisms also permit longer-range RBP interactions on mRNAs, in principle.

The studies on specific mRNAs clearly demonstrated the benefits of integration of multiple proteins into the regulation of transcripts. Combinations of cooperative or antagonistic interactions provide increased flexibility in the expression of a given mRNA depending on cellular context. Such combinatorial effects can also serve as safeguards to misregulation. In addition, these studies also highlighted the need for global, systematic, and unbiased methods to profile the extent of the regulatory RBP-RBP interactions.

## Transcriptome-wide integrative approaches to detect and characterize interactions

In recent years, global approaches have provided novel insights into the scope and complexity of combinatorial post-transcriptional regulation. These studies utilize various mixtures of computational and experimental procedures (Figure 1.3A), as described below.

**Analyses that rely on motif prediction for one or both RBPs**

Previously known RBP binding preference information, used in the form of sequence motifs to infer binding sites, alone or together with additional experimental data, has been informative in detecting interactions. The predicted sites for two or more RBPs can be analyzed across all 3' UTRs to determine proximity or extent of overlap, which suggests a co-regulatory relationship. For example, putative PUM1/2 and AU-rich sites were found to be enriched near predicted miRNA sites [74,77]. In some studies, transcripts harboring motifs were correlated to changes in mRNA expression upon RBP perturbation in order to

define the effects of RBP binding[76,78,79]. Alternatively, putative sites of proximal RBP-miRNA binding have been correlated with increased decay rates for the containing mRNAs, indicating joint regulation[74]. Additionally, RNA secondary structure predictions can integrate information about site accessibility or base-pairing in order to provide additional mechanistic insight. For example, in Incarnato et al.[76] overlap between PUM and miRNA motifs were found to occur at sites of low accessibility, and further evidence supported a model where Pumilio binding leads to an increase in accessibility of nearby miRNA sites. Similarly, miRNA seeds that are enriched near PUM sites showed complementarity to the PUM motif, which can be presumably disrupted by RBP binding[74]. It should be noted that since these studies are not supported by *in vivo* binding data, the results must be further validated to confirm the extent and nature of the interactions.

As experimental methods advanced, it became possible to analyze static *in vivo* binding datasets (CLIP-seq, RIP-seq, RIP-Chip) for one RBP with respect to nearby sequence features in order to determine potential co-regulation between RBPs[77,80]. These analyses are likely more sensitive in detecting interactions within the studied cell type, because in vivo binding data is used for one of the RBPs. In addition, the experimental datasets can also provide information regarding relative strength of RBP binding, which can be indicative of regulatory potential, and can be used to stratify or bin the transcripts in the analysis of motif co-occurence. Transcript abundance is then measured in wildtype and RBP perturbed cells to determine the regulatory effect of an RBP. This data can also be used to compare changes in transcript abundance for population predicted to be occupied by one RBP or co-occupied by both, which can provide insight into regulation. As before,

these studies require further validation, since informatically detected proximity or overlap in itself does not prove interaction.

**Analysis of static and dynamic in vivo binding datasets**

Integrative analysis of two (or more) CLIP sets from a single cellular condition or tissue has been increasingly applied in order to examine co-occupancy or overlap of sites. Here, the use of a second CLIP set (rather than sequence features) provides experimental evidence of binding location and strength within the studied cell type. In some cases, binding data is compared to a static RNA expression dataset[81]. More often, binding data is analyzed together with differential RNA-sequencing data upon RBP knockdown[75,82,83]. This allows for correlations between RBP binding and changes in RNA expression to be made, also taking into account the distance between sites. It is important to note that secondary effects of RBP knockdown are not ruled out in this type of analysis, so one cannot fully confirm that changes in transcript abundance are caused by direct binding of the RBP of interest. Rather, these studies provide evidence for regulatory action on a transcriptome-wide scale and can determine how the presence of multiple RBPs affects expression outcomes. For example, in HafezQorani et al.[75], transcripts that are bound by both HuR and MSI1 at sites within 200 nucleotides (nts) of each other showed a greater change HuR knockdown, compared to transcripts bound by either protein alone, indicating an interaction between the RBPs. Transcripts bound by both proteins further than 200 nts apart showed weaker combinatorial effects. This result was confirmed in HEK293 and Hela cells, and similar co-regulatory effects were also observed between HNRNBC and miR-148b.

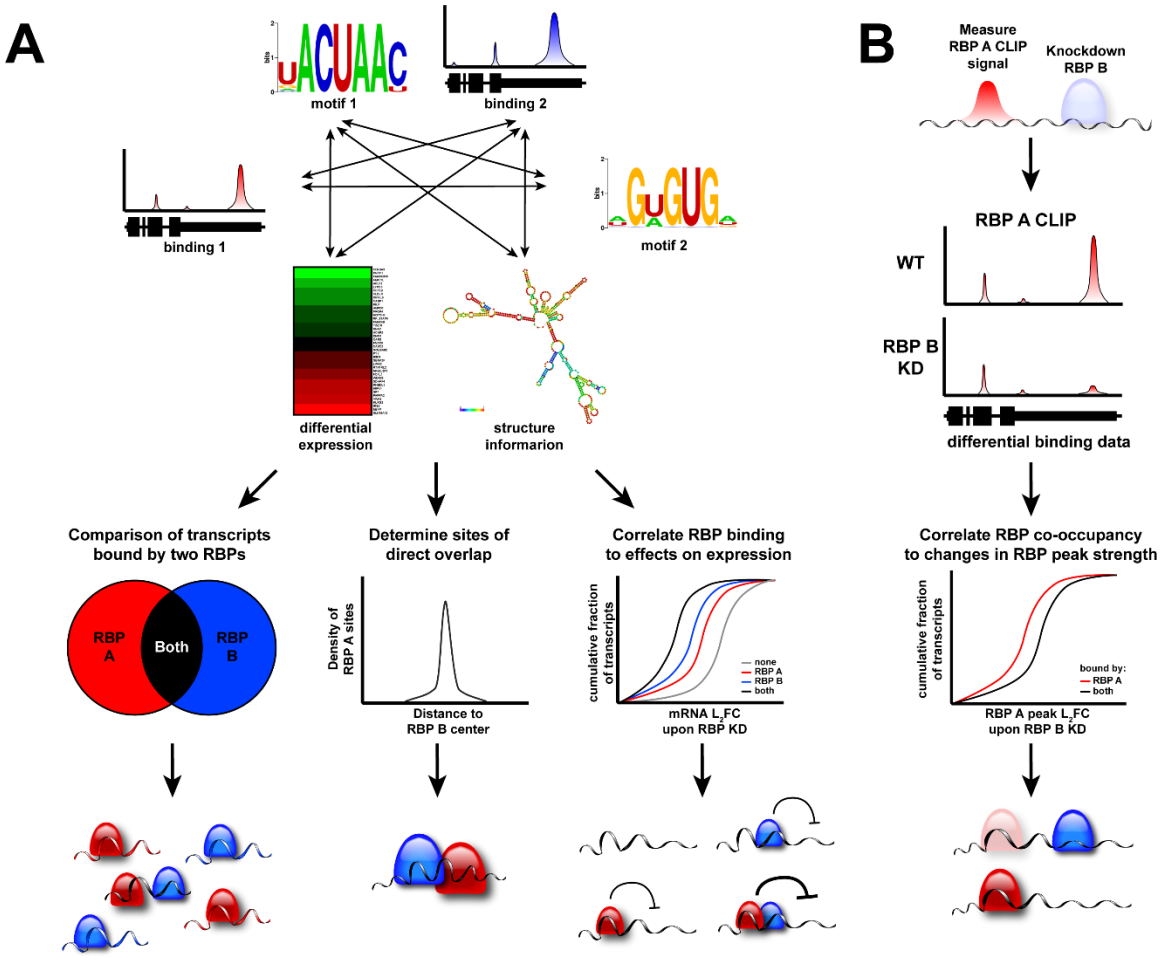



**Figure 1.3: Global interaction studies and outcomes.** (A) Possible combinations of computational and experimental data used to study RBP interactions, with downstream analysis and conclusions used to determine the extent of interaction between two RBPs. (B) Differential CLIP experimental design, downstream analysis, and conclusions. Red=RBP A, blue=RBP B.

To detect effects on co-occupancy more directly, CLIP-seq datasets for one RBP can be analyzed for differential binding under conditions of knockdown or knockout of its partner RBP[84-86] (Figure 1.3B). By quantifying changes in CLIP peak signal, this approach asks whether the presence of one RBP affects the binding of its partner across the transcriptome, rather than correlating RBP site proximity to downstream effects, such as changes in mRNA levels. However, because the observed CLIP intensity depends on underlying

mRNA abundance, it is important to normalize changes in CLIP signal to changes in transcript expression. To distinguish the extent of direct interactions from secondary effects of knockdown on a transcriptome-wide scale, changes in RBP1 peak strength upon RBP2 knockdown can be compared between transcripts occupied by both RBPs, versus transcripts bound by RBP1alone[84,85]. These analyses provide transcriptome-wide evidence of interactions. As the quality and reproducibility of CLIP-seq data improves, it should be possible to use statistical frameworks similar to RNA-seq to make inferences of RBP1 differential binding as a function of RBP2 presence on a per-site basis, providing much greater resolution. Further validation of functional combinatorial effects at such specific sites will be necessary.

Differential CLIP studies have made it clear that interactions between RBPs are highly context-dependent and employ different mechanisms. Two RBPs can act either cooperatively or antagonistically depending on the specific transcript, indicating that these interactions are affected by their local environment. Examples of this have been observed in co-regulation between the miRNA machinery and Pumilio[84], HuR[85], and PTB [87]. Because of these dynamic interactions, it will be important to consider the effects of multiple binding factors on a per transcript basis in order to get a clear regulatory picture. The transcriptome-wide binding approaches reinforce the view that most transcripts are under the regulatory control of multiple factors at any given time and building an integrated understanding of these interactions will allow for better predictive models.

## Meta-analyses support higher order RBP interaction networks

In addition to complex regulation due to the combinatorial nature of RBP interactions on individual transcripts, RBP expression levels across tissues and binding sites across mRNAs have evolved to co-regulate cellular processes by coordinately interacting with large groups of functionally related transcripts. Evidence for such interaction networks has been uncovered in large meta analyses, where many global datasets characterizing RBP binding, RBP localization, transcript abundance, and transcript function are mined to define patterns of regulation. As one may expect, the types of cellular processes an RBP is involved was found to correspond to the cellular localization of the RBP[37,88]. For example, nuclear RBPs encompass those involved in splicing, while cytoplasmic RBPs include those involved in the regulation of translation. Similarly, the location of RBP binding on a transcript was shown to predict its function: RBPs that are involved in common regulatory processes tend to co-localize to the same general region (5' or 3' UTR, CDS, splice sites)[88]. In addition, groups of RBPs displayed overlap in regulating functionally related sets of transcripts, with distinct mechanisms with respect to their synthesis, processing, translation and degradation rates[37,88,89]. Detailed integrative analysis of multiple CLIP-seq datasets also identified the presence of 3' UTR regulatory hotspots, , narrow locations that can recruit many RBPs[82,89]. These hotspots drastically increase the potential for combinatorial interactions and challenge us to consider mechanistic models that involve co-regulation between more than two RBPs at a time.

Interactions between RBPs can also be hierarchical, where a subset of RBPs are responsible for the regulation of other RBP transcripts, which then would cause broader downstream

effects. In one study, RBPs were categorized into “clusters” which connected cooperative and antagonistic interactions between RBPs of a transcript and “chains” which included RBPs that regulated clusters and thus were capable of coordinating a broad set of biological processes[90].

These studies support the existence of “RNA operons”, as previously proposed[91]. Analogous to the classic bacterial DNA operon, and RNA operon model allows for the coordinated expression of many transcripts in order to execute a particular cellular function. In the simplest version of this model, the expression of one RBP leads to changes in expression of a set of transcripts that are functionally related, allowing for activation or repression of a given process. Combinatorial interactions between RBPs add a layer of complexity to the model by allowing a greater amount of fine-tuning and a larger selection of cellular outcomes.

## Phase separation drives large scale organization of RNP particles

In addition and in juxtaposition to the site- and RBP-specific interaction models described above, cellular RNAs and RBPs participate in less specific and more fluid and multivalent interactions that result in aggregation of large numbers of RBPs. Such phase-separated droplets containing high densities of RNP particles have become a focus of research in recent years[14,92,93]. The droplets are thought to be formed through weak associations between the intrinsically disordered regions (IDRs) within many RBPs, as well as through additional specific and non-specific RBP-RNA and RNA-RNA interactions. Altogether, these interactions bring many RNP components together into a membrane-less organelle.

RBPs have individual propensities for homo- and hetero-association and can form distinct assemblies that may have different functional biological roles. However, the instability of IDR interactions allows droplets to be highly dynamic and reversible. As a result, these droplets can quickly change the localization, concentration, and availability of RNP particles within the cell, providing spatial and temporal control. Although the role of many of these droplets is still under investigation, they are known to have biological importance. The nucleolus is one example, where high concentrations of factors responsible for ribosomal synthesis are brought together in order to increase efficiency[93,94]. Misregulation of the fluidity of these droplets is observed in neurodegenerative disorders, where droplet interactions become irreversibly stable and drive the formation of solid aggregates, or plaques[95-99].

The overall regulatory interactions between RBPs and RNA therefore span a spectrum of mechanisms. On the one side are a multitude of weak and low-specificity interactions, which drive the dynamics of phase separation. On the other are stronger, more binary interactions between RBPs and RNAs, which drive most of the binding at specific 3' UTR sites and impart distinct regulatory fates on particular messages, as discussed in this review. However, it is the combination of these two types of interactions that comprise the full cellular response. Large condensates can increase local concentrations of RBPs, which in turn facilitates recruitment of RBP binding to its target transcripts. Conversely, these condensates may act to filter sets of RBPs and RNAs, which in turn can separate a given RBP from its target RNA, preventing its regulatory effects.

## Conclusions

RBPs execute a wide range of functions that are important in almost all aspects of cellular biology and understanding regulatory effects of RBPs has been an important area of study. Although recent studies have shed significant light on how individual RBPs function, many RBPs remain uncharacterized. Further research into these RBPs is essential for a clearer understanding of cellular processes as well as mechanisms of disease.

Recent advances in the study of combinatorial effects of RBPs have revealed the interconnected nature of RBP regulation. The RNA landscape is a crowded environment, making combinatorial interactions between RBPs common, and a greater understanding of the complexity of this landscape is necessary. In addition, RBP interactions are context-specific and influenced by the local environment of their target transcript. These context-specific changes impact the regulatory behavior of the RBP, which in turn can dramatically change outcomes of gene expression. Further, RBPs coordinate regulation not only on a single transcript, but across many transcripts in order to modulate cellular processes. The coming years still have much to decipher regarding how RBP interactions are coordinated and to develop more comprehensive and powerful regulatory models.

## References


1 Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15**, 829-845, doi:10.1038/nrg3813 (2014).

2 Muller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nature Reviews Genetics* **14**, 275-287, doi:10.1038/nrg3434 (2013).

3 Beckmann, B. M. *et al.* The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat. Commun.* **6**, 10127, doi:10.1038/ncomms10127 (2015).

4 Cook, K. B., Hughes, T. R. & Morris, Q. D. High-throughput characterization of protein-RNA interactions. *Brief. Funct. Genomics* **14**, 74-89, doi:10.1093/bfgp/elu047 (2015).

5 Cook, K. B., Kazan, H., Zuberi, K., Morris, Q. & Hughes, T. R. RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.* **39**, D301-308, doi:10.1093/nar/gkq1069 (2011).

6 Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* **8**, 479-490, doi:10.1038/nrm2178 (2007).

7 Ray, D. *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172-177, doi:10.1038/nature12311 (2013).

8 Djuranovic, S., Nahvi, A. & Green, R. A parsimonious model for gene regulation by miRNAs. *Science* **331**, 550-553, doi:10.1126/science.1191138 (2011).

9 Fabian, M. R., Sonenberg, N. & Filipowicz, W. Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* **79**, 351-379, doi:10.1146/annurev-biochem-060308-103103 (2010).

10 Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20-51, doi:10.1016/j.cell.2018.03.006 (2018).

11 Colegrove-Otero, L. J., Minshall, N. & Standart, N. RNA-binding proteins in early development. *Crit. Rev. Biochem. Mol. Biol.* **40**, 21-73, doi:10.1080/10409230590918612 (2005).

12 Castello, A., Fischer, B., Hentze, M. W. & Preiss, T. RNA-binding proteins in Mendelian disease. *Trends Genet.* **29**, 318-327, doi:https://doi.org/10.1016/j.tig.2013.01.004 (2013).

13 Brinegar, A. E. & Cooper, T. A. Roles for RNA-binding proteins in development and disease. *Brain Res.* **1647**, 1-8, doi:10.1016/j.brainres.2016.02.050 (2016).

14 Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, doi:10.1126/science.aaf4382 (2017).

15 Degrauwe, N., Suva, M. L., Janiszewska, M., Riggi, N. & Stamenkovic, I. IMPs: an RNA-binding protein family that provides a link between stem cell maintenance in normal development and cancer. *Genes Dev.* **30**, 2459-2474, doi:10.1101/gad.287540.116 (2016).

16 Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. & Gilad, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**, 1509-1517, doi:10.1101/gr.079558.108 (2008).
17 Mann, M. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **7**, 952-958, doi:10.1038/nrm2067 (2006).
18 Faoro, C. & Ataide, S. F. Ribonomic approaches to study the RNA-binding proteome. *FEBS Lett.* **588**, 3649-3664, doi:10.1016/j.febslet.2014.07.039 (2014).
19 Lin, J. S. & Lai, E. M. Protein-Protein Interactions: Co-Immunoprecipitation. *Methods Mol. Biol.* **1615**, 211-219, doi:10.1007/978-1-4939-7033-9_17 (2017).
20 Wan, Y. *et al.* Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**, 706-709, doi:10.1038/nature12946 (2014).
21 Bellaousov, S., Reuter, J. S., Seetin, M. G. & Mathews, D. H. RNAstructure: Web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res.* **41**, W471-474, doi:10.1093/nar/gkt290 (2013).
22 Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E. & Weeks, K. M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959-965, doi:10.1038/nmeth.3029 (2014).
23 Ke, S. *et al.* A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* **29**, 2037-2053, doi:10.1101/gad.269415.115 (2015).
24 Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206, doi:10.1038/nature11112 (2012).
25 Ray, D. *et al.* RNAcompete methodology and application to determine sequence preferences of unconventional RNA-binding proteins. *Methods* **118-119**, 3-15, doi:10.1016/j.ymeth.2016.12.003 (2017).
26 Lambert, N. *et al.* RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell* **54**, 887-900, doi:10.1016/j.molcel.2014.04.016 (2014).
27 Keene, J. D., Komisarow, J. M. & Friedersdorf, M. B. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat. Protoc.* **1**, 302-307, doi:10.1038/nprot.2006.47 (2006).
28 Zambelli, F. & Pavesi, G. RIP-Seq data analysis to determine RNA-protein associations. *Methods Mol. Biol.* **1269**, 293-303, doi:10.1007/978-1-4939-2291-8_18 (2015).
29 Lu, Z., Guan, X., Schmidt, C. A. & Matera, A. G. RIP-seq analysis of eukaryotic Sm proteins identifies three major categories of Sm-containing ribonucleoproteins. *Genome Biol.* **15**, R7, doi:10.1186/gb-2014-15-1-r7 (2014).
30 Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464-469, doi:10.1038/nature07488 (2008).
31 Zhang, C. & Darnell, R. B. Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* **29**, 607-614, doi:10.1038/nbt.1873 (2011).

32 Ule, J., Jensen, K., Mele, A. & Darnell, R. B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**, 376-386, doi:10.1016/j.ymeth.2005.07.018 (2005).
33 Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129-141, doi:10.1016/j.cell.2010.03.009 (2010).
34 Konig, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17**, 909-915, doi:10.1038/nsmb.1838 (2010).
35 Van Nostrand, E. L. *et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* **13**, 508-514, doi:10.1038/nmeth.3810 (2016).
36 Bailey, T. L. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653-1659, doi:10.1093/bioinformatics/btr261 (2011).
37 Van Nostrand, E. L. *et al.* A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. 179648, doi:10.1101/179648 %J bioRxiv (2018).
38 Iadevaia, V. & Gerber, A. P. Combinatorial Control of mRNA Fates by RNA-Binding Proteins and Non-Coding RNAs. *Biomolecules* **5**, 2207-2222, doi:10.3390/biom5042207 (2015).
39 Dassi, E. Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins. *Frontiers in molecular biosciences* **4**, 67, doi:10.3389/fmolb.2017.00067 (2017).
40 Jiang, P. & Coller, H. Functional interactions between microRNAs and RNA binding proteins. *Microrna* **1**, 70-79 (2012).
41 Kedde, M. *et al.* A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol.* **12**, 1014-1020, doi:10.1038/ncb2105 (2010).
42 Miles, W. O., Tschop, K., Herr, A., Ji, J. Y. & Dyson, N. J. Pumilio facilitates miRNA regulation of the E2F3 oncogene. *Genes Dev.* **26**, 356-368, doi:10.1101/gad.182568.111 (2012).
43 Kim, H. H. *et al.* HuR recruits let-7/RISC to repress c-Myc expression. *Genes Dev.* **23**, 1743-1748, doi:10.1101/gad.1812509 (2009).
44 Glorian, V. *et al.* HuR-dependent loading of miRNA RISC to the mRNA encoding the Ras-related small GTPase RhoB controls its translation during UV-induced apoptosis. *Cell Death Differ.* **18**, 1692-1701, doi:10.1038/cdd.2011.35 (2011).
45 Nakamura, H. *et al.* Cooperative role of the RNA-binding proteins Hzf and HuR in p53 activation. *Mol. Cell. Biol.* **31**, 1997-2009, doi:10.1128/MCB.01424-10 (2011).
46 Jing, Q. *et al.* Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell* **120**, 623-634, doi:10.1016/j.cell.2004.12.038 (2005).
47 Rahman, M. A. *et al.* HnRNP L and hnRNP LL antagonistically modulate PTB-mediated splicing suppression of CHRNA1 pre-mRNA. *Sci. Rep.* **3**, 2931, doi:10.1038/srep02931 (2013).

48 Damianov, A. *et al.* Rbfox Proteins Regulate Splicing as Part of a Large Multiprotein Complex LASR. *Cell* **165**, 606-619, doi:10.1016/j.cell.2016.03.040 (2016).
49 Coelho, M. B. *et al.* Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. *EMBO J.* **34**, 653-668, doi:10.15252/embj.201489852 (2015).
50 Okunola, H. L. & Krainer, A. R. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Mol. Cell. Biol.* **29**, 5620-5631, doi:10.1128/mcb.01678-08 (2009).
51 Bradley, T., Cook, M. E. & Blanchette, M. SR proteins control a complex network of RNA-processing events. *RNA* **21**, 75-92, doi:10.1261/rna.043893.113 (2015).
52 Weill, L., Belloc, E., Castellazzi, C. L. & Mendez, R. Musashi 1 regulates the timing and extent of meiotic mRNA translational activation by promoting the use of specific CPEs. *Nat. Struct. Mol. Biol.* **24**, 672-681, doi:10.1038/nsmb.3434 (2017).
53 Peng, Y., Yuan, J., Zhang, Z. & Chang, X. Cytoplasmic poly(A)-binding protein 1 (PABPC1) interacts with the RNA-binding protein hnRNPLL and thereby regulates immunoglobulin secretion in plasma cells. *J. Biol. Chem.* **292**, 12285-12295, doi:10.1074/jbc.M117.794834 (2017).
54 Reznik, B., Clement, S. L. & Lykke-Andersen, J. hnRNP F complexes with tristetraprolin and stimulates ARE-mRNA decay. *PLoS One* **9**, e100992, doi:10.1371/journal.pone.0100992 (2014).
55 Min, K. W. *et al.* AUF1 facilitates microRNA-mediated gene silencing. *Nucleic Acids Res.* **45**, 6064-6073, doi:10.1093/nar/gkx149 (2017).
56 Weidensdorfer, D. *et al.* Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA* **15**, 104-115, doi:10.1261/rna.1175909 (2009).
57 Anantharaman, A. *et al.* RNA-editing enzymes ADAR1 and ADAR2 coordinately regulate the editing and expression of Ctn RNA. *FEBS Lett.* **591**, 2890-2904, doi:10.1002/1873-3468.12795 (2017).
58 Copsey, A. C. *et al.* The helicase, DDX3X, interacts with poly(A)-binding protein 1 (PABP1) and caprin-1 at the leading edge of migrating fibroblasts and is required for efficient cell spreading. *Biochem. J.* **474**, 3109-3120, doi:10.1042/bcj20170354 (2017).
59 Kedde, M. *et al.* RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell* **131**, 1273-1286, doi:10.1016/j.cell.2007.11.034 (2007).
60 Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I. & Filipowicz, W. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125**, 1111-1124, doi:10.1016/j.cell.2006.04.031 (2006).
61 Kundu, P., Fabian, M. R., Sonenberg, N., Bhattacharyya, S. N. & Filipowicz, W. HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. *Nucleic Acids Res.* **40**, 5088-5100, doi:10.1093/nar/gks148 (2012).

62 Liu, L. *et al.* Competition between RNA-binding proteins CELF1 and HuR modulates MYC translation and intestinal epithelium renewal. *Mol. Biol. Cell* **26**, 1797-1810, doi:10.1091/mbc.E14-11-1500 (2015).
63 Topisirovic, I. *et al.* Stability of eukaryotic translation initiation factor 4E mRNA is regulated by HuR, and this activity is dysregulated in cancer. *Mol. Cell. Biol.* **29**, 1152-1162, doi:10.1128/MCB.01532-08 (2009).
64 Tiedje, C. *et al.* The p38/MK2-driven exchange between tristetraprolin and HuR regulates AU-rich element-dependent translation. *PLoS Genet.* **8**, e1002977, doi:10.1371/journal.pgen.1002977 (2012).
65 Ho, J. J. D. *et al.* Active stabilization of human endothelial nitric oxide synthase mRNA by hnRNP E1 protects against antisense RNA and microRNAs. *Mol. Cell. Biol.* **33**, 2029-2046, doi:10.1128/MCB.01257-12 (2013).
66 Zarnack, K. *et al.* Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* **152**, 453-466, doi:10.1016/j.cell.2012.12.023 (2013).
67 Zhu, J., Mayeda, A. & Krainer, A. R. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell* **8**, 1351-1361 (2001).
68 Gazzara, M. R. *et al.* Ancient antagonism between CELF and RBFOX families tunes mRNA splicing outcomes. *Genome Res.* **27**, 1360-1370, doi:10.1101/gr.220517.117 (2017).
69 Hall, M. P. *et al.* Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation. *RNA* **19**, 627-638, doi:10.1261/rna.038422.113 (2013).
70 Lu, G. & Hall, T. M. T. Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure* **19**, 361-367, doi:10.1016/j.str.2010.12.019 (2011).
71 Kim, S. H., Shanware, N. P., Bowler, M. J. & Tibbetts, R. S. Amyotrophic lateral sclerosis-associated proteins TDP-43 and FUS/TLS function in a common biochemical complex to co-regulate HDAC6 mRNA. *J. Biol. Chem.* **285**, 34097-34105, doi:10.1074/jbc.M110.154831 (2010).
72 Chen, C. A. *et al.* Antagonistic actions of two human Pan3 isoforms on global mRNA turnover. *RNA* **23**, 1404-1418, doi:10.1261/rna.061556.117 (2017).
73 Hopkins, T. G. *et al.* The RNA-binding protein LARP1 is a post-transcriptional regulator of survival and tumorigenesis in ovarian cancer. *Nucleic Acids Res.* **44**, 1227-1246, doi:10.1093/nar/gkv1515 (2016).
74 Jiang, P., Singh, M. & Coller, H. A. Computational assessment of the cooperativity between RNA binding proteins and MicroRNAs in Transcript Decay. *PLoS Comput. Biol.* **9**, e1003075, doi:10.1371/journal.pcbi.1003075 (2013).
75 HafezQorani, S. *et al.* Modeling the combined effect of RNA-binding proteins and microRNAs in post-transcriptional regulation. *Nucleic Acids Res.* **44**, e83, doi:10.1093/nar/gkw048 (2016).
76 Incarnato, D., Neri, F., Diamanti, D. & Oliviero, S. MREdictor: a two-step dynamic interaction model that accounts for mRNA accessibility and Pumilio

binding accurately predicts microRNA targets. *Nucleic Acids Res.* **41**, 8421-8433, doi:10.1093/nar/gkt629 (2013).
77 Galgano, A. *et al.* Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. *PLoS One* **3**, e3164, doi:10.1371/journal.pone.0003164 (2008).
78 Leibovich, L., Mandel-Gutfreund, Y. & Yakhini, Z. A structural-based statistical approach suggests a cooperative activity of PUM1 and miR-410 in human 3'-untranslated regions. *Silence* **1**, 17, doi:10.1186/1758-907X-1-17 (2010).
79 Jacobsen, A., Wen, J., Marks, D. S. & Krogh, A. Signatures of RNA binding proteins globally coupled to effective microRNA target sites. *Genome Res.* **20**, 1010-1019, doi:10.1101/gr.103259.109 (2010).
80 Lebedeva, S. *et al.* Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol. Cell* **43**, 340-352, doi:10.1016/j.molcel.2011.06.008 (2011).
81 Mukherjee, N. *et al.* Global target mRNA specification and regulation by the RNA-binding protein ZFP36. *Genome Biol.* **15**, R12, doi:10.1186/gb-2014-15-1-r12 (2014).
82 Plass, M., Rasmussen, S. H. & Krogh, A. Highly accessible AU-rich regions in 3' untranslated regions are hotspots for binding of regulatory factors. *PLoS Comput. Biol.* **13**, e1005460, doi:10.1371/journal.pcbi.1005460 (2017).
83 Mukherjee, N. *et al.* Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol. Cell* **43**, 327-339, doi:10.1016/j.molcel.2011.06.007 (2011).
84 Sternburg, E. L., Estep, J. A., Nguyen, D. K., Li, Y. & Karginov, F. V. Antagonistic and cooperative AGO2-PUM interactions in regulating mRNAs. *Sci. Rep.* **8**, 15316, doi:10.1038/s41598-018-33596-4 (2018).
85 Li, Y., Estep, J. A. & Karginov, F. V. Transcriptome-wide Identification and Validation of Interactions between the miRNA Machinery and HuR on mRNA Targets. *J. Mol. Biol.* **430**, 285-296, doi:10.1016/j.jmb.2017.12.006 (2018).
86 Lu, Y. C. *et al.* ELAVL1 modulates transcriptome-wide miRNA binding in murine macrophages. *Cell Rep.* **9**, 2330-2343, doi:10.1016/j.celrep.2014.11.030 (2014).
87 Xue, Y. *et al.* Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell* **152**, 82-96, doi:10.1016/j.cell.2012.11.045 (2013).
88 Mukherjee, N. *et al.* Deciphering human ribonucleoprotein regulatory networks. *Nucleic Acids Res.* **47**, 570-581, doi:10.1093/nar/gky1185 (2019).
89 Li, Y. E. *et al.* Identification of high-confidence RNA regulatory elements by combinatorial classification of RNA-protein binding sites. *Genome Biol.* **18**, 169, doi:10.1186/s13059-017-1298-8 (2017).
90 Quattrone, A. & Dassi, E. The architecture of the human RNA-binding protein regulatory network. 041426, doi:10.1101/041426 %J bioRxiv (2018).
91 Keene, J. D. RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* **8**, 533-543, doi:10.1038/nrg2111 (2007).

92 Weber, S. C. & Brangwynne, C. P. Getting RNA and protein in phase. *Cell* **149**, 1188-1191, doi:10.1016/j.cell.2012.05.022 (2012).
93 Zhu, L. & Brangwynne, C. P. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* **34**, 23-30, doi:10.1016/j.ceb.2015.04.003 (2015).
94 Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686-1697, doi:10.1016/j.cell.2016.04.047 (2016).
95 Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133, doi:10.1126/science.1134108 (2006).
96 Kwiatkowski, T. J., Jr. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205-1208, doi:10.1126/science.1166066 (2009).
97 Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208-1211, doi:10.1126/science.1165942 (2009).
98 Vanderweyde, T. *et al.* Interaction of tau with the RNA-Binding Protein TIA1 Regulates tau Pathophysiology and Toxicity. *Cell Rep.* **15**, 1455-1466, doi:10.1016/j.celrep.2016.04.045 (2016).
99 Vanderweyde, T., Youmans, K., Liu-Yesucevitz, L. & Wolozin, B. Role of stress granules and RNA-binding proteins in neurodegeneration: a mini-review. *Gerontology* **59**, 524-533, doi:10.1159/000354170 (2013).

# CHAPTER 2

## Antagonistic and Cooperative AGO2-PUM Interactions in Regulating mRNAs.


## Abstract

Approximately 1500 RNA-binding proteins (RBPs) profoundly impact mammalian cellular function by controlling distinct sets of transcripts, often using sequence-specific binding to 3′ untranslated regions (UTRs) to regulate mRNA stability and translation. Aside from their individual effects, higher-order combinatorial interactions between RBPs on specific mRNAs have been proposed to underpin the regulatory network. To assess the extent of such co-regulatory control, we took a global experimental approach followed by targeted validation to examine interactions between two well-characterized and highly conserved RBPs, Argonaute2 (AGO2) and Pumilio (PUM1 and PUM2). Transcriptome-wide changes in AGO2-mRNA binding upon PUM knockdown were quantified by CLIP-seq, and the presence of PUM binding on the same 3′ UTR corresponded with cooperative and antagonistic effects on AGO2 occupancy. In addition, PUM binding sites that overlap with AGO2 showed differential, weakened binding profiles upon abrogation of AGO2 association, indicative of cooperative interactions. In luciferase reporter validation of candidate 3′ UTR sites where AGO2 and PUM colocalized, three sites were identified to host antagonistic interactions, where PUM counteracts miRNA-guided repression. Interestingly, the binding sites for the two proteins are too far for potential antagonism due to steric hindrance, suggesting an alternate mechanism. Our data experimentally confirm the combinatorial regulatory model and indicate that the mostly repressive PUM proteins

can change their behavior in a context-dependent manner. Overall, the approach underscores the importance of further elucidation of complex interactions between RBPs and their transcriptome-wide extent.

## Introduction

Post-transcriptional control of gene expression is central to a wide range of cellular processes, ensuring proper cell homeostasis. In addition, it allows for rapid alterations in gene expression, which are necessary for correct developmental transitions, as well as response to environmental changes. The regulation is mainly accomplished by RNA-binding proteins (RBPs) and microRNAs (miRNAs), both of which target mature messenger RNAs by binding to defined sites in the 3′ untranslated region (UTR). Often, these specific factors serve as the target recognition components of larger complexes, which recruit additional, nonspecific factors to stabilize or repress target mRNA expression. Furthermore, the presence of multiple regulatory complexes is thought to have a combinatorial effect on the mRNA. However, while much is known about the individual effects of RBPs and microRNAs, how interactions between the complexes affect downstream gene expression remains to be fully understood.

As a major part of this regulatory model, miRNAs in stable association with Argonaute (AGO) proteins recruit repressive complexes to target mRNA sites through perfect complementarity within the miRNA "seed" (nucleotides 2-8) and imperfect base-pairing throughout the rest of the 22-23 nt miRNA[1-3]. The Argonaute-containing RNA-induced silencing complex (RISC) engages decapping/deadenylation enzymes to cause mRNA destabilization or leads to translational inhibition by a mechanism that is still under investigation[4-6]. MicroRNA-guided silencing is expansive across the transcriptome, regulating more than 60% of all genes at different developmental and physiological states[7].

Pumilio proteins are a prototypical group of predominantly repressive sequence-specific mRNA regulators which are conserved across eukaryotes and participate in similar regulatory processes in many species[8-10]. The family is classified by the Pumilio Homology domain (Pum-HD), which adopts an arc-shaped fold capable of directly binding RNA[10-12]. The canonical Pum-HD is composed of eight alpha-helical repeats, each recognizing one nucleotide of its binding motif, 5′-UGUAnAUA[13,14]. In *Drosophila,* Pumilio proteins regulate embryonic development[15-18], neuronal function[19-21], and germline development/maintenance, with the latter function also observed in *C. elegans*[9,22-25].

In mammals, there are two Pumilio proteins, PUM1 and PUM2, which carry out similar functions with some redundancy and specificity. The paralogs are nearly identical in the Pum-HD (89% in human), have an apparently indistinguishable RNA binding motif[13,26-28], and share a fairly ubiquitous profile of expression[29,30]. Pum1,2 double knockout mice are inviable[27], while single knockouts demonstrate roles in spermatogenesis and primordial folliculogenesis[31-33]. Similarly, individual roles for Pum2[34-36] and Pum1[37] in mammalian neuronal function have been shown, while defects of the neural-specific double knockout are more profound[27]. Consistent with their expression patterns, the proteins are likely to have functions in broader tissues, as cell-based assays show roles in genome stability and cell cycle regulation[38-40]. In reporter assays, the two proteins demonstrate functional redundancy, since depletion of both factors is necessary to abrogate their repressive effects[39,41]. However, Pumilio proteins appear to bind substantially distinct sets of transcripts *in vivo*[26,27].

Pumilio proteins can function as repressive factors through mRNA destabilization and translational inhibition. Messenger RNA decay is accomplished by recruiting deadenylation/decapping complexes and poly(A) binding protein (PABP) antagonism[41-43], resulting in a decrease of target mRNA levels[13,38,44]. Although the full mechanism of translational inhibition is still unknown, interference with translation elongation or termination, and competition with eIF4E have been proposed[45-47]. However, Pumilio proteins have also been documented to have stabilizing effects on target mRNAs across eukaryotes[34,44,48-51]. The mechanisms underlying these biologically relevant effects appear to be case-specific and may involve interaction with additional factors and/or changes in 3′ UTR secondary structure upon Pumilio binding.

Previous studies provide several examples of interactions between Argonaute and Pumilio in altering gene expression. The two proteins have been shown to act cooperatively to regulate the CDKN1B mRNA, where PUM binding increases AGO binding by causing an mRNA secondary structure switch[39]. A cooperative interaction is also observed on the E2F3 transcript, where PUM binding leads to increased AGO recruitment. Here, alternative polyadenylation can eliminate PUM sites in the 3′ UTR, a strategy utilized by cancer cells to escape cell cycle regulation[40]. In addition, transcriptome-wide studies have identified evolutionary and functional evidence of interaction: it has been shown that predicted miRNA binding sites are enriched in the vicinity of PUM sites, and that co-occurrence of PUM and miRNA sites in stem loops or sites of low accessibility correlates with repression of the mRNA[26,52-55]. However, the global extent of the impact of PUM on AGO2 binding and vice versa, as well as the functional consequences of such interactions, has not been

experimentally examined with specific site resolution. In the present study, the RBPs' mRNA occupancy profiles were quantified by CLIP-seq as a function of their partner's presence, identifying both cooperative and antagonistic effects of UTR co-occupancy and site overlap on binding. Additionally, the data defined differences and similarities in PUM1 and PUM2 interactions with AGO2, providing information on the paralogs' roles. Finally, individual validation experiments confirmed co-regulatory expression control by PUM and AGO for a subset of identified sites.

## Results

### Determination of AGO2, PUM1 and PUM2 binding profiles and their inter-dependencies by CLIP-seq

The individual binding behavior of Ago and Pum homologs on mammalian mRNAs has been examined, but their effect on each other's binding and function is largely undetermined transcriptome-wide. Since the human Argonaute homologs AGO1-3 associate with similar sets of miRNAs[56,57], mRNAs[28], and proteins[58], and AGO4 is typically poorly expressed, our analysis focused on the more abundant AGO2 homolog. We have previously used a reciprocal perturbation approach to assess the global AGO2-HuR interactions by quantitative CLIP-seq[59]. To further investigate potential AGO2-PUM co-regulatory interactions, crosslinking and immunoprecipitation followed by sequencing (CLIP-seq) experiments were performed for the endogenous PUM1, PUM2 and AGO2 proteins in 293 cells and derivatives (Appendix A). For each protein, binding profiles were quantified under multiple conditions. To determine if PUM proteins impact AGO2 binding behavior, AGO2 CLIP-seq was carried out in conditions of PUM1 or PUM2 knockdown with two distinct siRNAs each (average 82-91% knockdown, Supplementary Figure 2.1), and compared to control siRNA knockdowns. Three biological replicates of PUM1 KD and four replicates of PUM2 KD were performed, resulting in 2,187,553 and 2,565,487 PCR-collapsed read counts, respectively. Conversely, the impact of AGO association with mRNAs on PUM binding was assessed by PUM1 and PUM2 CLIP-seq in wildtype cells and two independent DICER-deficient clones[60]. Cells lacking DICER cannot produce mature canonical miRNAs and are used to abolish miRNA-guided targeting of the four

AGO proteins to mRNAs. After sequencing, filtering and processing, 3028 AGO2, 9727 PUM1 and 4988 PUM2 sites in 3′ UTRs were identified in these datasets. To verify the accuracy of the binding sites, comparisons were made to a previously reported PUM2 PAR-CLIP dataset[28] and several AGO2 and AGO1 HITS-CLIP, PAR-CLIP and CLASH datasets[28,61-63] obtained from starBase[64]. The number of overlapping peaks with each set were computed. Similarly, the overlaps with each set in 100 control experiments, where the positions of peaks in the data were randomized on the same UTR, were used to calculate the mean and standard deviation of the background overlap expectation. Z scores of the actual overlap numbers were calculated to be 11.9 - 33.9 for PUM1 and PUM2 overlaps with previous data, and ranged from 6.4 to 51.9 for AGO2 overlaps, indicating very significant agreement between the datasets. Additionally, searches for miRNA seed site complements and PUM sequence motifs were performed (see below).

**Distinct and similar binding characteristics of PUM1 and PUM2**

Human PUM1 and PUM2 share 69% identity / 74% similarity along their entire length, and their mouse orthologs have partially redundant but distinct functions in regulating the cell cycle and developmental processes[27,31,33]. However, their transcriptome-wide binding repertoires have not been examined and contrasted in detail. To this end, we compared the CLIP-sequencing datasets collected for PUM1 and PUM2. At the level of transcripts, PUM2 bound to 2969 3′ UTRs, a large majority of which, 2154, were also bound by PUM1, consistent with a partially redundant role for PUM2. PUM1 occupied a broader set of transcripts, with an approximately equal number of 3′ UTRs not bound by PUM2 (Figure 2.1A). *De novo* motif enrichment analysis[65] in sites bound by PUM1 and PUM2 (Figure

2.1A, DREME E-values of $5.6e^{-200}$ and $2.5e^{-64}$, respectively) revealed versions of the previously determined PUM motif as the top identified sequences[10-12].

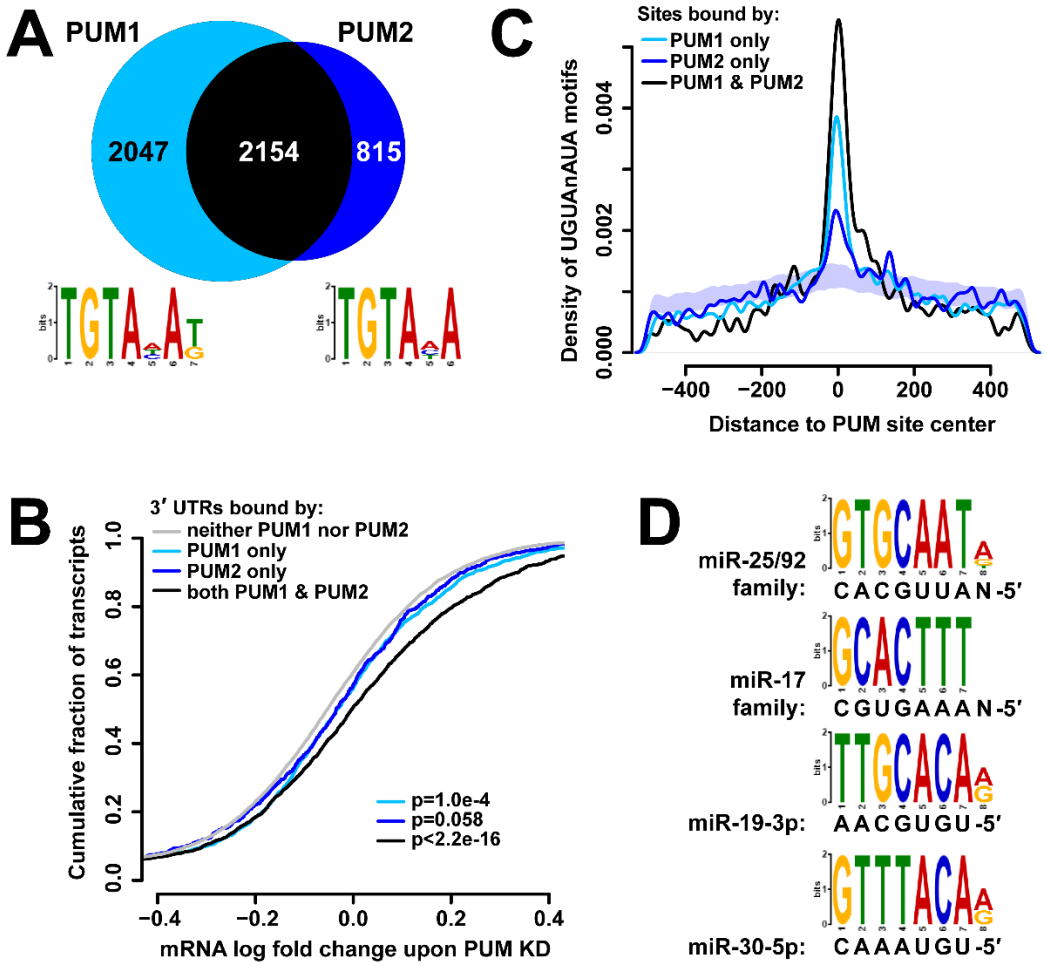



**Figure 2.1: PUM1 and PUM2 have distinct and similar binding characteristics. A)** Number of transcripts with 3′ UTRs bound by PUM1 (light blue), PUM2 (dark blue), or both (black). Motif logos at the bottom depict the top-scoring enriched motif identified by DREME for the PUM1 and PUM2 sites, consistent with the previously determined PUM motif. **B)** Cumulative distribution plot of mRNA level log fold change upon PUM knockdown[52] for populations of transcripts with 3′ UTRs bound by PUM1 (light blue), PUM2 (dark blue), both (black), or neither (gray). p-values correspond to a Kolmogorov-Smirnov test of each transcript set against the rest of the transcriptome. Cumulative distribution plot of mRNA level log fold change upon PUM knockdown for populations of transcripts with 3′ UTRs not bound by AGO2 is shown in Supplementary Figure 2.13. **C)** Density of PUM motifs near PUM1 (light blue), PUM2 (dark blue), or overlapping (black) PUM sites. The shaded area represents a mean +/-1 SD interval of 100 control densities where PUM motif locations were randomized within each 3′ UTR, and distances to PUM2-only peaks computed. **D)** DREME differential motif enrichment between PUM2 and PUM1 bound sequences show an enrichment of miRNA seed complements within the PUM2 sites.

To assess the effects of PUM binding on mRNA levels, we compared the distributions of mRNA abundance log fold changes upon knockdown of both PUM1 and PUM2[52] for transcript classes with various combinations of PUM1 and PUM2 binding in the 3′ UTR (Figure 2.1B). Transcripts bound by PUM1 or PUM2 alone showed stabilization of mRNA levels, albeit below statistical significance for PUM2 ($p = 0.058$, Kolmogorov-Smirnov (KS) test). This suggests that such uniquely bound populations of transcripts experience changes in abundance, supporting a non-redundant function at least for PUM1. However, transcripts bound by both PUM1 and PUM2 showed a substantially larger increase in abundance upon PUM knockdown, indicating that jointly targeted transcripts are under more regulation. Furthermore, the extent of PUM1 and PUM2 CLIP signal on 3′ UTRs correlated with changes in mRNA levels upon PUM KD (Supplementary Figure 2.2), aided by the elimination of PCR bias in read counts through the use of random barcodes in the ligated CLIP adaptor[66]. Similar observations have been made for the correlation of AGO CLIP strength with target repression[67], indicating that CLIP data can quantitatively reflect downstream effects on the transcripts.

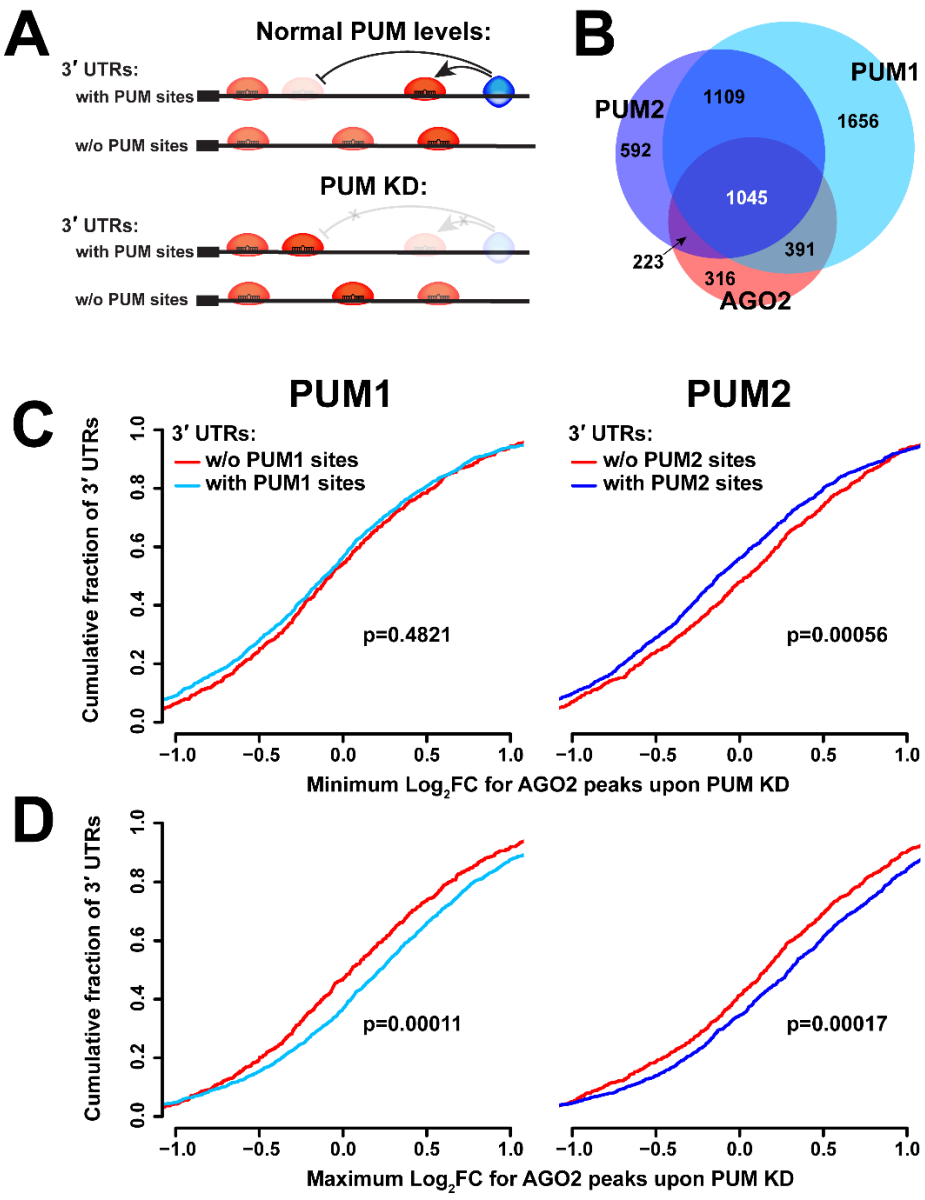



**Figure 2.2: PUM and AGO2 occupy similar transcript populations, and AGO2 binding is affected by PUM presence on the same UTR**. **A)** Schematic of potential interactions between AGO2 and the PUM proteins. **B)** Number of 3′ UTRs bound by any combination of PUM1 (light blue), PUM2 (dark blue), and AGO2 (red). **C)** Cumulative distribution plot of minimum log fold change for AGO2 peaks on 3′ UTRs with and without Pumilio proteins. **D)** Cumulative distribution plot of maximum log fold change for AGO2 peaks on 3′ UTRs with and without Pumilio proteins.

At the level of individual binding sites, the two proteins exhibited an even lower extent of overlap, with 1226 of 9727 PUM1 peaks and 1233 of 4988 PUM2 peaks overlapping each other (Supplementary Figure 2.3). The PUM1- and PUM2-only populations of peaks were still enriched in the canonical PUM motif sequences (Figure 2.1C), suggesting that some of the identified paralog-only sites possess functionality and binding specificity. Interestingly, a distinguishing sequence feature emerges from differential motif enrichment between PUM1 and PUM2 binding sites: sequences matching the seed complements of four abundant miRNAs are enriched in PUM2 relative to PUM1 sites (Figure 2.1D). Such an enrichment suggests PUM2 and AGO binding sites are often found in close proximity and have the potential for interaction.

**AGO2 binding is affected by PUM presence on the same 3′ UTR, suggesting direct interactions**

AGO and PUM proteins have been shown to interact cooperatively on the CDKN1B and E2F3 transcripts[39,40], and previous transcriptome-wide analyses indicate that predicted miRNA seeds and PUM motifs in 3′ UTRs are enriched in each other's vicinity, often in self-complementary secondary structures, and such arrangements lead to faster transcript decay[26,52,54,68]. To further understand interactions between AGO and PUM, the extent of AGO2 binding at 3′ UTRs (quantified by CLIP-seq) were compared between control and PUM knockdown conditions and normalized to the changes in mRNA levels upon PUM1 and PUM2 knockdown reported in[52]. Our initial analysis aimed to determine whether AGO2 binding is influenced by the presence or absence of PUM on the same 3′ UTR

(Figure 2.2A). A majority of transcripts bound by AGO2 were also occupied by one or both of the PUM proteins (Figure 2.2B), indicating a large potential for interactions.

Each UTR can possess multiple AGO2 sites with individual responses to the presence of PUM (i.e. log fold changes in CLIP-seq signal upon PUM KD after mRNA level correction; LFCs), and a minority of all sites are expected to interact with PUM. Thus, to examine transcriptome-wide evidence of co-regulation, we hypothesized that the AGO2 sites with the greatest changes for each transcript (the sites with the minimal LFC and maximal LFC within each UTR) are the most likely candidates for interaction, and the analysis focused on such sites. UTRs were then pooled into two categories for comparison: those bound by both AGO2 and PUM1 or PUM2 (therefore potentially interacting on the 3′ UTR per se), and a control group bound only by AGO2 and not the corresponding PUM, thus incapable of an interaction on the UTR, (i.e. representing secondary, background effects). When examining AGO2 sites with the minimal LFC upon PUM2 knockdown within each UTR, sites on transcripts co-occupied by PUM2 showed lower minimal LFCs, compared to peaks on transcripts bound by AGO2 alone (Figure 2.2C). This difference suggests the presence of cooperative binding interactions between PUM2 and AGO2, since co-occupancy with PUM2 corresponds with more AGO2 dissociation upon PUM2 KD. Interestingly, such a relationship was not observed between PUM1 and AGO2. Conversely, the maximum log fold change of AGO2 binding upon PUM knockdown is higher on transcripts that are co-occupied by PUM proteins (stronger association), compared to transcripts bound by AGO2 alone (Figure 2.2D), consistent with AGO-PUM antagonism. In summary, these results indicate that AGO2 binding on PUM-bound 3′ UTRs is

differentially impacted by PUM knockdown, strongly suggesting interactions between the proteins on the same 3′ UTR.

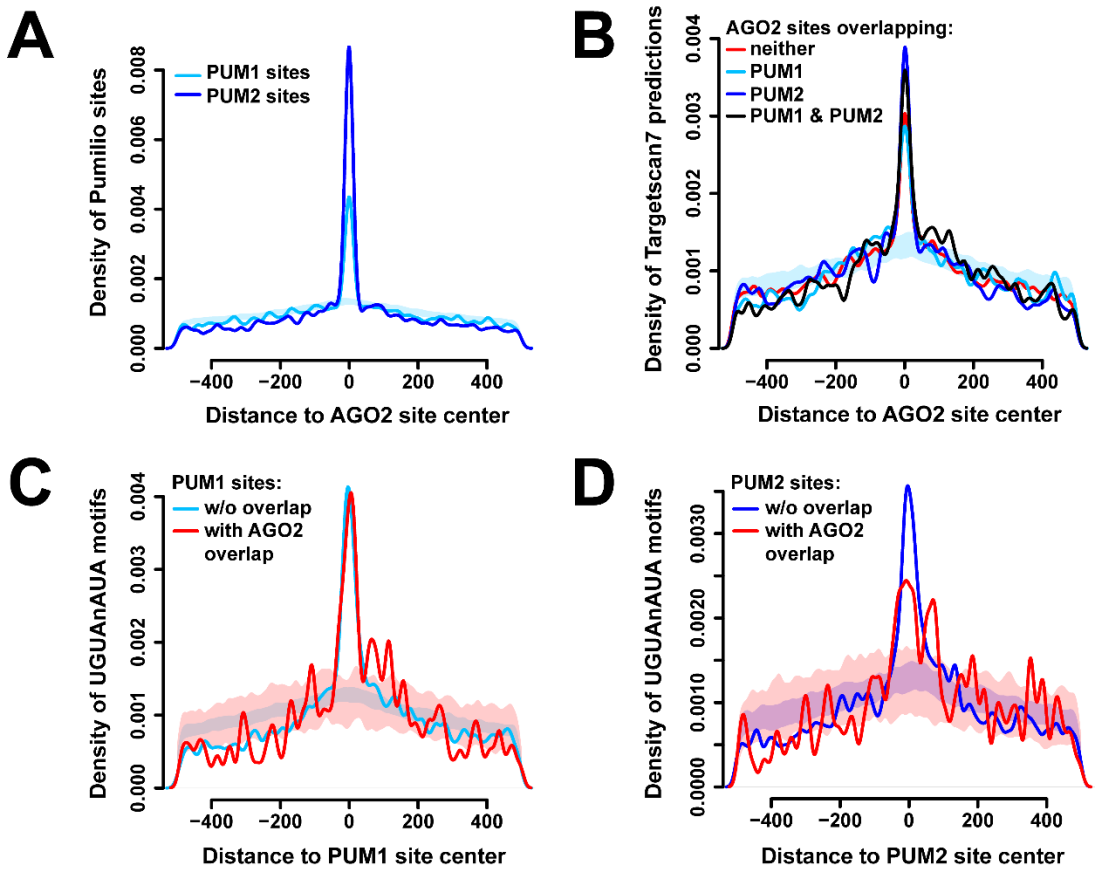



**Figure 2.3: AGO2 and PUM proteins show overlap at the binding site level. A)** Density of PUM1 (light blue) and PUM2 (dark blue) binding sites surrounding AGO2 binding sites. The shaded area represents a mean +/-1 SD interval of 100 control densities where PUM1 site locations were randomized within each 3′ UTR, and distances to AGO2 peaks computed. **B)** Density of Targetscan7 predictions surrounding AGO2 peaks, separated by sites that overlap with PUM1 (light blue), PUM2 (dark blue), both (black), or neither (red). The shaded area represents randomized PUM1-only controls as above. **C, D)** Density of PUM motifs surrounding PUM1 (C) and PUM2 (D) sites, separated by populations that do (red) or do not (blue) overlap with AGO2. The shaded areas represent randomized controls for the corresponding populations.

**AGO2-PUM site co-occupancy affects PUM binding**

Next, the extent of overlap between AGO2 and PUM sites, and its potential effects on protein binding, were analyzed. Of the 3028 AGO2 sites across all 3′ UTRs, a significant fraction - 794 and 800 sites - overlapped with PUM1 or PUM2, respectively (Supplementary Figure 2.4). Notably, the number of AGO2/PUM2 overlaps is larger, despite the roughly two-fold fewer total PUM2 vs. PUM1 sites. The extent of co-occupancy is also evident in positional enrichment of PUM1 and PUM2 around AGO2 sites (Figure 2.3A), with PUM2 again being more dominant. These results are consistent with the differential enrichment of miRNA seed complements observed in PUM2 sites relative to PUM1 (Figure 2.1D). To ensure that the overlapping peaks represent true AGO2 and PUM populations and not procedural CLIP artifacts (such as cross-linking hot spots), the density of Targetscan7-predicted miRNA seed complements[69] within AGO2 peaks were examined, which demonstrated enrichment regardless of whether they overlap with the PUM proteins (Figure 2.3B). Additionally, the experimentally determined AGO2 sites were found to be centered on the predicted miRNA seed occurrences (Supplementary Figure 2.5). Similarly, PUM binding sites show enrichment of the consensus motif with and without AGO2 overlap (Figure 2.3C and 2.3D). To assess if specific miRNA families are used in targeting individual vs. PUM-overlapping AGO2 sites, occurrences of each miRNA family seed complement in sites were tallied and compared. Overlapping and non-overlapping AGO2 sites were found to have similar miRNA repertoires (Spearman correlation coefficient of 0.651 and 0.647 for PUM1 and PUM2 overlap, respectively), suggesting that AGO2 sites with PUM overlap do not prefer specific miRNAs.

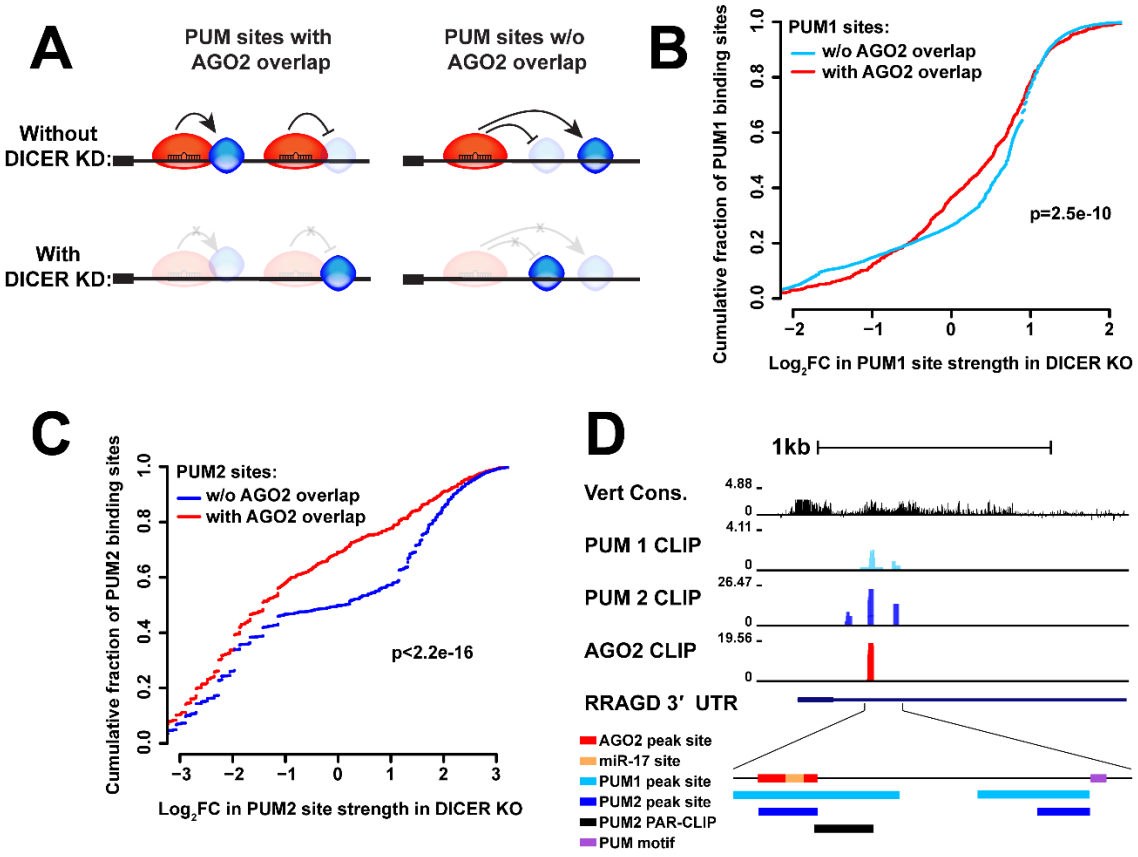



**Figure 2.4: AGO2-PUM site co-occupancy affects PUM binding. A)** Schematic of possible interactions between AGO2 and the PUM proteins for PUM sites with and without AGO2 overlap, as well as predicted effect of DICER knockdown in each case. **B)** Cumulative distribution plots of the log fold change in PUM1 site strength, with (red) and without (light blue) AGO2 overlap. **C)** Cumulative distribution plots of the log fold change in PUM2 site strength, with (red) and without (dark blue) AGO2 overlap. **D)** Example UCSC browser view of a site with overlapping PUM1 (light blue), PUM2 (dark blue), and AGO2 (red) binding within the RRAGD 3′ UTR. PUM1 and PUM2 binding sites show overlap with an independently generated PUM2 PAR-CLIP and are in close proximity to a PUM motif. The AGO2 binding site contains a predicted miR-17 site.

To determine whether overlap with AGO2 affects PUM occupancy, we examined the change in PUM1 and PUM2 binding strength between wildtype and DICER-deficient cells, where miRNA-guided AGO2 association with transcripts is abolished[60]. Comparing the LFCs of PUM binding upon DICER loss among two groups of PUM sites (those with and without AGO2 overlap, Figure 2.4A) revealed that both PUM1 and PUM2 binding are differentially weakened at sites of co-occupancy (Figure 2.4B and 2.4C). In this analysis,

PUM2 also exhibited a stronger dependence on AGO2 association than PUM1. Thus, the loss of PUM binding as a result of eliminated AGO2 binding suggests the presence and predominance of cooperative (as opposed to antagonistic) interactions between the two proteins.

**Validation of CLIP-seq candidate sites demonstrates antagonistic PUM-AGO2 interactions**

Since the transcriptome-wide analysis suggested that overlapping sites may host interactions, candidates were selected to test whether co-regulatory effects on protein expression could be detected. To increase the likelihood of identifying such sites, we picked from the top 20 strongest normalized and unnormalized AGO2 binding sites based on overlap with a PUM2 binding site in our data and with a previously published PUM2 PAR-CLIP dataset[28], yielding 26 sites. An example UCSC browser shot of a candidate site is shown in Figure 2.4D. Sites ranging 76-330 nt in length, including the above elements with 10 nt of flanking sequence, were 4X concatenated and cloned into the 3′ UTR of the *Renilla* luciferase reporter gene in the psiCheck2 plasmid with an internal firefly normalization control. In parallel, constructs with a mutated (shuffled) AGO2 site were also generated. The reporters were used to determine the regulatory activity of the AGO site, as the ratio between wildtype and mutant luciferase activity (WT/mut ratio), in three distinct cellular settings. Wildtype T-REx-293 cells reported on the site's activity in the presence of PUM, and DICER-deficient 293T derivatives[60] uncovered the miRNA-targeted contribution to the regulation. In turn, to assess the effect of PUM on the regulation, PUM double knockout (PDKO, Supplementary Figure 2.6) cells were generated using CRISPR-

Cas9 combined with homologous recombination[70]. To confirm that PUM has been functionally knocked out, a Renilla luciferase reporter with a 4X concatenated sequence of a known Pumilio-regulated site from the 3′ UTR of the *Drosophila* hunchback mRNA was used (Figure 2.5A). Similarly, as a control for miRNA-guided regulation, a strong AGO2 site in the LRIG3 3′ UTR with a predicted miRNA seed that does not contain a PUM motif was chosen and cloned as a 4X concatamer. Confirming the expectation that the AGO2 site is repressive, the WT/mut ratio was shown to be significantly less than one in T-Rex-293 cells, and the repression was DICER- but not PUM-dependent (Figure 2.5B).

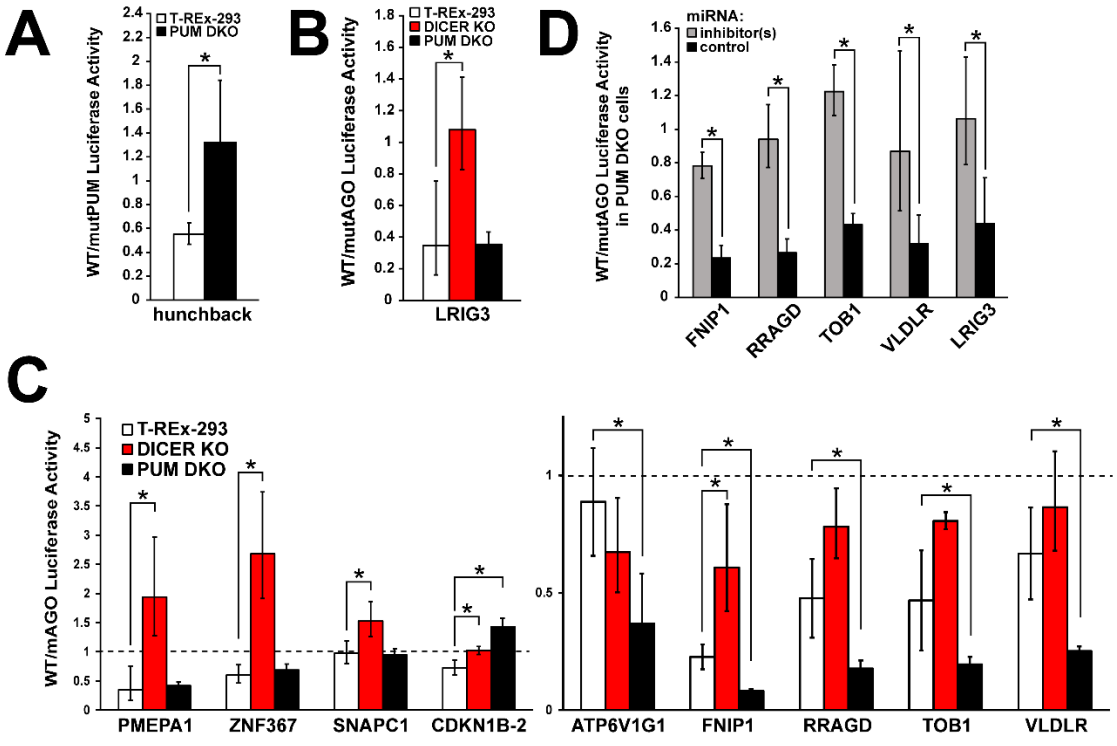



**Figure 2.5: Luciferase reporter assays of candidate sites reveal antagonistic AGO2-PUM interactions.** **A)** A known Pumilio-regulated site from the 3′ UTR of the Drosophila hunchback mRNA shows loss of regulation in PUM double knockout cells. **B)** A strong AGO2 site in the LRIG3 3′ UTR with a predicted miRNA seed shows loss of repression in DICER knockout, but not PUM double knockout, cells. **C)** A subset of candidates show miRNA-dependent activity (PMEPA1, ZNF367, SNAPC1) or both miRNA and PUM dependence (CDKN1B-2, ATP6V1G1, FNIP1, RRAGD, TOB1, VLDLR). Candidate sites were tested in three conditions: wildtype T-REx-293, DICER knockout cells, and PUM double knockout cells. **D)** Sites co-

regulated by AGO and PUM sites show dependence on specific miRNAs in PUM double knockout cells. For FNIP1 and VLDLR, WT and mutant AGO peak constructs were used. For RRAGD and TOB1, WT and mutant miRNA seed constructs were used. *= $p<0.05$.

When tested in the above settings, the candidate panel demonstrated several regulatory patterns. Candidates that showed no regulation, a stabilizing/repressive site activity that was not dependent on either PUM or DICER, or excessively high protective site activity, (Supplementary Figure 2.7) were not examined further. Two sites (PMEPA1 and ZNF367, Figure 2.5C) possessed a DICER-dependent, but PUM-independent repressive activity, representing canonical AGO regulation. The SNAPC1 site showed no regulation in WT and PDKO cells, but stabilization in DICER KO cells, suggesting a miRNA dependence that does not fit the canonical model, or indirect effects (Figure 2.5C). Interestingly, our selection criteria, informed by precise AGO2 and PUM binding locations, identified two overlapping sites in the CDKN1B 3′ UTR, where a PUM/AGO2 interaction mediated by two miR-221/222 seed sites was previously demonstrated[39]. For the CDKN1B-1 construct, which included one of the previously implicated miR-221/222 seeds under the AGO2 site (Supplementary Figure 2.8), mutation of the site did not reveal any statistically significant differences in regulation between WT and PUM- or DICER-deficient cells (Supplementary Figure 2.8). However, activity at a separate, 3′-proximal CDKN1B-2 site was dependent on both DICER and PUM.

Finally, five sites demonstrated either no effect in WT cells, or a repression that was attenuated in DICER KOs; interestingly, all five became substantially more repressive in PDKOs (Figure 2.5C, ATP6V1G1, FNIP1, RRAGD, TOB1, VLDLR). This pattern suggests a previously unobserved antagonistic interaction between PUM and AGO. Since

the repressive activity of the AGO site in wildtype cells was minimal for most of the sites, we wanted to confirm whether the increased repression in PDKO cells was directed by the miRNA machinery. To this end, inhibition of specific target miRNAs was performed in PDKO cells for four of the five antagonistic sites, and ratios of WT to AGO site mutant reporter constructs were measured as above (Figure 2.5D). ATP6V1G1 was excluded from further analysis because it did not contain a predicted miRNA seed. The increase in repression in the absence of PUM was confirmed to be miRNA dependent (Figure 2.5D), indicating that miRNA-guided repression can occur at these sites, but is prevented when Pumilio proteins are present in the cell.

**PUM antagonizes AGO through the predicted Pumilio motif**

Next, we aimed to determine whether the observed antagonism is due to the presence of PUM binding to the predicted motif(s) adjacent to the AGO site, or if this change in AGO activity is due to secondary or trans-regulatory effects. Constructs with mutations in the PUM motifs (perfect and imperfect) instead of the AGO site were generated. miRNA dependence was again determined by comparing luciferase activity under specific and control miRNA inhibitor conditions, while the effects of PUM on the miRNA-dependent activity were then tested under wildtype site, mutant PUM site, and PDKO conditions. Importantly, VLDLR and RRAGD showed equivalent repression whether the protein, its sites, or both were removed (Figure 2.6A), indicating that regulation by PUM is only due to PUM's action at the predicted sites. Activity of the FNIP1 and TOB1 reporters in this assay was inconsistent with such a model (Supplementary Figure 2.9): PUM effects for TOB1 were not significant, while FNIP1 showed increased repression for the mutant site

construct that did not reach statistical significance, and a further, significant repression when the protein was removed (Figure 2.6A). However, removal of both the PUM protein and site restored FNIP1 miRNA repression to WT levels. One interpretation is that the mutated PUM site also serves as a binding site for a third RBP that cooperatively interacts with AGO2. Overall, the results for VLDLR and RRAGD demonstrate an antagonistic model of interaction between AGO and PUM at overlapping sites.

To determine whether a physical AGO2-PUM interaction could be detected, co-immunoprecipitation experiments were performed. Immunoprecipitation (IP) of AGO2 followed by immunoblot of the two PUM proteins did not show any interaction (Supplementary Figure 2.10A). However, IP of both PUM1 and PUM2 showed a weak interaction with AGO2. For PUM1, additional washes disrupted most of this interaction, and RNase treatment abolished the interaction entirely (Supplementary Figure 2.10B). For PUM2, a weak interaction was detected in the IP (Supplementary Figure 2.10C), but the PUM2-bead complex was lost in subsequent treatments. These results suggest that AGO2 and PUM proteins do not form a stable stoichiometric complex and may enter into relatively weak and transient interactions. Consistent with these observations, binding of endogenous PUM2 to tagged AGO2 has also been detected by IP-mass spectrometry in one study[71], but was not identified in another[58].

Finally, in order to confirm that results determined in artificial constructs correspond to effects on endogenous messages. RT-qPCR was performed to determine mRNA levels of both RRAGD and VLDLR mRNAs in wildtype and PDKO cells. Results show that both transcripts are downregulated in PDKO cells when compared to WT (Figure 2.6B). This

result is consistent with regulation observed in the reporter constructs, and demonstrates an unexpected, protective activity in conjunction with AGO2, in contrast to Pumilio's canonical repressive role.

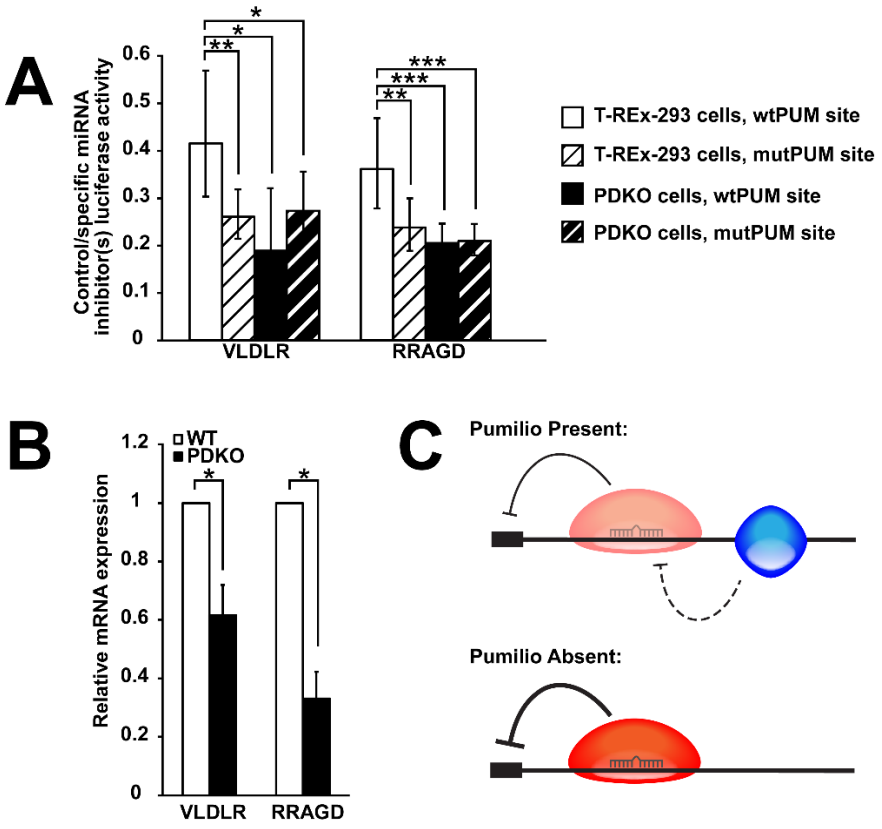



**Figure 2.6: PUM antagonizes AGO through the predicted PUM motif, and PUM double knockout stabilizes endogenous transcripts. A)** Sites co-regulated by AGO and PUM show dependencies on specific miRNAs, PUM proteins, and the predicted PUM site. Luciferase reporter constructs co-transfected with control or specific miRNA inhibitors were tested under wildtype (white) and PUM double knockout (black) conditions with WT (solid) or mutant PUM sites (stripes). **B)** RT-qPCR of endogenous VLDLR and RRAGD transcripts in both wildtype (white) and PUM double knockout (black) cells. **C)** A model of the antagonistic effect of PUM on AGO2 regulation of VLDLR and RRAGD mRNAs. *= $p<0.05$, **=$p<0.01$, ***= $p<0.001$

## Discussion

We employed a transcriptome-wide approach to determine the binding profiles and inter-dependencies of AGO2, PUM1, and PUM2 sites on mRNA 3′ UTRs. In agreement with the extensive homology between the two PUM proteins, binding sites of both paralogs showed enrichment of a nearly identical motif consistent with the known UGUAnAUA recognition sequence, and there was a significant amount of overlap in the populations of transcripts bound by the proteins. However, PUM1 bound a substantially larger population of transcripts than PUM2 (Figure 2.1A). Interestingly, previous studies utilizing a different method (RIP-Chip) or different antibodies in CLIP-seq similarly identified approximately twice as many transcripts associated with PUM1 vs PUM2, with comparable overlap[26,27], strongly suggesting that these observations reflect their intrinsic binding properties. The greater number of transcripts may be explained by a higher binding affinity (possibly mediated by PUM1's longer N-terminal extension), or additional specificity determinants associated with PUM2 binding. Transcripts bound exclusively by PUM1 showed a detectable amount of repression, suggesting that PUM1-specific targeting of mRNA populations is biologically relevant (Figure 2.1B). Nevertheless, transcripts bound by both paralogs showed substantially greater regulation. Overlap of individual PUM1 and PUM2 binding sites was significantly less common than that of 3′ UTRs, with paralog-specific loci still enriched in the binding motif (Figure 2.1C). Such loci included peaks from abundant transcripts supported by many individually barcoded ligation event read counts (Supplementary Figure 2.3) indicating that the observed specificity is unlikely to arise from undersampling. It is yet to be determined how PUM proteins can occupy distinct binding

sites despite a shared motif and extensive protein homology, although potential interaction partners may be involved. In accord with this model, differential motif discovery between PUM1 and PUM2 sites showed an enrichment of miRNA seeds near PUM2 sites (Figure 2.1D), suggesting a greater extent of interaction between PUM2 and the miRNA machinery.

Quantitative examination of AGO2 and PUM CLIP data while manipulating the levels/occupancy of the other factor provided direct transcriptome-wide evidence of interactions between the RBPs at the binding level. The majority of AGO2-bound transcripts were also occupied by one or both of the PUM proteins (Figure 2.2B), and changes in PUM protein level lead to changes in AGO2 binding strength on co-occupied 3′ UTRs. Here, PUM2 showed a signature of cooperative binding with AGO2, and both factors demonstrated antagonistic interactions (Figure 2.2C, D). Further, AGO2 showed significant overlap with the PUM proteins at the binding site level, and PUM binding at such loci was differentially weakened upon elimination of AGO2 association, suggesting an overall predominance of cooperative interactions between the factors. Again, PUM2 demonstrated a stronger effect in this analysis (Figure 2.4C) and had a higher degree of overlap with AGO2 sites (Figure 2.3A). It should be noted that AGO2 CLIP data was collected under independent PUM1 or PUM2 knockdowns, and not simultaneous reduction of both paralogs. This design allows for a better understanding of the individual interactions of the two Pumilios with AGO2, which are not well studied on a transcriptome-wide scale. However, this design also likely leads to smaller or undetectable interaction signal in cases where the two paralogs have mostly overlapping and redundant roles.

Although the analysis focused on the canonical 3′ UTR regulatory region, AGO2, PUM1, and PUM2 sites and their overlap were also observed in other annotation categories. Consistent with the size and function of the coding sequence (CDS) and the 5′ UTR, the number of sites for all three proteins was lower in these mRNA regions compared to the 3′ UTR, while overlap between AGO and the PUM proteins, was similar or lower (Supplementary Figure 2.4A, B). However, enrichment of the PUM consensus motif was only weakly detected in PUM1 (but not PUM2) CDS sites, (Supplementary Figure 2.4C), and was undetectable in the 5′ UTR (data not shown). Consistent with motif enrichment for PUM1 sites in the CDS and its interactions with AGO2 in the 3′ UTR (Figure 2.4B), LFCs of PUM occupancy upon DICER loss showed that PUM1 binding is differentially weakened at CDS sites of AGO2 co-occupancy (Supplementary Figure 2.11B). At sites where motif enrichment was not found (PUM2 in the CDS, or PUM1,2 in the 5′ UTR), no such effects were observed (Supplementary Figure 2.11). These conclusions are consistent with previously reported observations that PUM motifs in CDS, but not 5′ UTR regions are enriched[26] and correlate with mRNA repression by PUM[44]. Additionally, substantial AGO2-PUM site overlap, but no PUM motif enrichment, was detected in the ncRNA category, potentially driven by many spurious interactions of the RBPs with abundant cellular ncRNAs.

Validation of individual overlapping sites using luciferase reporters demonstrated a cooperative interaction on the CDKN1B mRNA, a known target of AGO2-PUM co-regulation[39], and unexpectedly identified two novel instances where PUM can antagonize AGO at a nearby site (Figure 2.5C). For CDKN1B, a 5′-proximal site that partially overlaps

the previously characterized miR-221/222 seeds did not exhibit co-regulation in our assay, potentially because further necessary nearby sequence elements were not included in the construct. However, a separate, 3′-proximal overlapping site showed both a PUM and miRNA dependence, consistent with a cooperative interaction, suggesting that AGO2-PUM co-regulation at this 3′ UTR occurs through more than one site (Figure 2.5C). Lack of a predicted miRNA seed at the second site prevented further examination of its regulation. The identification of antagonistic co-regulation of VLDLR and RRAGD mRNAs prompts a model where binding of PUM near the AGO2 sites leads to attenuation of miRNA-guided repression (Figure 2.6C), showing that Pumilio, a normally repressive factor, can take on a stabilizing role in a context-dependent manner, and underscores the flexibility that is imparted on the regulation by combinatorial interactions. These observations are consistent with other studies that have suggested a stabilizing role for the Pumilio proteins[44,49,72,73] and suggest that further cases of stabilization by PUM may occur through the antagonism of other regulatory RBPs. While the precise mechanism(s) of AGO-PUM antagonism are not understood, involvement of secondary structure rearrangements is a good starting hypothesis that has been previously implicated[39,52,55]. Alternatively, steric clashes may be responsible, although in the VLDLR and RRAGD constructs the PUM motifs and AGO binding sites (narrowed down to the predicted miRNA seeds within) appear to be far enough away from each other that the individual factors would not interfere physically. However, since both proteins recruit large complexes, interference may still be possible. It is possible that additional candidates from

the interrogated set involved AGO2-PUM co-regulation, but the required sequence elements were not fully included in the reporter constructs.

In many cases, candidate sites selected for luciferase validation contained multiple Pumilio motifs, including the full consensus sequence (UGUAnAUA) and shorter versions with degeneracies in the more weakly defined 3′ end (UGUAnnUA, UGUAnAnA, or UGUAnAUn). For example, the VLDLR site included three perfect motifs, and all three sites were mutated for testing. Individual site mutants would determine the contributions of each of the sites to the observed antagonistic effect. In contrast, the RRAGD site contained one full and one imperfect PUM motif, which together were sufficient for co-regulation.

While testing of individual targets revealed instances where PUM impacted AGO2 activity, AGO's reciprocal ability to affect PUM binding was also observed in the CLIP data (Figure 2.4B, C). However, the reporter experimental design was not amenable to identify such interactions, since PUM site mutants for the full list of candidates were not independently tested in wildtype and DICER-deficient conditions to isolate the PUM site activity and its dependence on miRNAs. Thus, further testing would likely uncover additional examples and a more diverse interaction profile.

The identified switches from repression to protection for PUM underscores the dynamics of RBP regulation and the necessity to develop a comprehensive, site-specific understanding of their effects on gene expression. For major regulatory proteins such as AGO and PUM, misinterpretation or incomplete information about available regulatory modes can hinder the mechanistic understanding of disease states. For example, previous

studies have shown that misregulation of VLDLR (very low density lipoprotein receptor) leads to many neurodevelopmental disorders, including cerebellar ataxia, mental retardation and disequilibrium syndrome (CAMRQ1) and cerebellar hypoplasia[74-76]. Understanding that this gene is a target for AGO/PUM co-regulation, and knowing that PUM acts to stabilize the transcript in this context, can potentially direct new therapeutic strategies. Similar arguments can be made for RRAGD (Ras-related GTP-binding protein D), which is a component within the amino acid sensing branch of mTORC1 signaling. RRAGD misregulation is correlated with renal and liver cancer prognosis[30]. A better understanding of how protein interactions affect the outcome of gene expression will lead to more precise disease treatment and fewer off-target effects.

Our results expand on a growing number of studies demonstrating interactions of the miRNA machinery on specific mRNAs with PUM and other RBPs, including HuR[77,78], SFPQ[79], PTB[80], and DND1[81], reviewed in[53,82], as well as global analyses that identify miRNA-RBP interactions[52,54,83,84]. Overall, the presented pairwise perturbation studies, together with similar efforts on other RBPs[59] and broader static analyses of RBP co-occupancy on 3′ UTRs[68] will be necessary to uncover the full extent of combinatorial post-transcriptional regulation of mRNAs.

## Materials and Methods

### Cell culture and PUM knockdown

T-REx-293 cells were obtained from Invitrogen, and DICER deficient cells (along with the parental line) were a kind gift from B. Cullen[60]. PUM double knockout cells were generated as previously described[70]. All cells were grown in DMEM (Corning) with 10% fetal bovine serum (Corning) and 10 units/ml of penicillin/streptomycin (Gibco) at 37 °C with 5% $CO_2$.

For PUM1 and PUM2 knockdowns, 3 and 4 biological replicates, respectively, of T-REx-293 cells were separately transfected with two distinct siRNAs against either PUM1 or PUM2, or with a GL3.1 siRNA control (Supplementary File 2). TransIT-TKO (Mirus) transfections with 100 nM siRNA were performed for the first PUM1 and PUM2 replicates, and calcium phosphate transfections with 100 nM siRNA were carried out for the later replicates. For each replicate set, three successive transfections 2-3 days apart were performed to get sufficient knockdown. For each replicate/condition, three to six 15-cm plates of cells were collected for the CLIP procedure.

### AGO2, PUM1, and PUM2 HITS-CLIP and data analysis

The AGO2 CLIP protocol was performed as previously described[85]. Mouse anti-AGO2 (Santa Cruz, clone 4F9) antibody was used for AGO2 CLIP and goat anti-PUM1 (Bethyl, A300-201A) and rabbit anti-PUM2 (Bethyl, A300-202A) antibodies were used for PUM1 and PUM2 CLIP, respectively. Two replicate sets of PUM1 CLIP and one set of PUM2 CLIP, were performed in 293T cells (control), NoDice (2–20) and NoDice (4–25) cells.

Libraries were sequenced on an Illumina HiSeq2500 instrument with a multiplex of six libraries per lane. Sequencing data was analyzed as previously described[59], with readcount cutoffs equal to the number of samples in each CLIP set.

**PUM double knockout cell generation**

PUM2 single KO cells were generated as described[70] and were used as the parental line for generating PUM double knockout cells. Single guide RNAs (sgRNAs) targeting exon 4 and exon 15 of the PUM1 gene were each cloned into the pLx330 Cas9-sgRNA expression plasmid (Supplementary File 2). A hygromycin resistance cassette flanked by two 900 nt homology regions within exon 4 and exon 15 were assembled in the pUC-19 vector as described[70] and used as the template for homology directed repair to replace most of the PUM1 gene with the resistance cassette. Cells were transfected with all three plasmids concurrently, and hygromycin selection was applied after three days. Clonal populations of cells were generated and screened using primers flanking the exon 4 CRISPR-Cas9 cut site (Supplementary File 2). Candidate clones were validated by western blot for absence of both PUM1 and PUM2 using antibodies described above.

**Reporter vector plasmid construction**

AGO2/PUM overlapping sites were selected based on criteria outlined in the text. For the purposes of selection of candidates among overlapping sites, AGO2 sites/peak locations and widths were defined from a combined dataset containing AGO2 CLIP data of PUM KDs of the current study, and HuR KDs of Li et al[59]. The top 20 sites sorted by unnormalized or mRNA-level-normalized AGO2 CLIP signal level shared many

candidates, and both sets were included in the selection. Candidate sites included the entire AGO2 and PUM peaks of interest, expanded to sequences corresponding to miRNA seeds and PUM motifs within 200 nt, plus 10 nt of flanking sequence. Sites were assembled into 4x concatamers into the 3′ UTR of the Renilla luciferase reporter gene contained in the psiCHECK-2 (Promega) plasmid as previously described[59]. In parallel, mutant constructs were generated where the entire AGO2 CLIP peak sequence was shuffled in order to abolish AGO2 binding. All WT and mutant monomer sequences along with primers used for assembly are listed in Supplementary File S3. Hunchback and LRIG3 positive controls were assembled with the same design.

For RRAGD, FNIP1, VLDLR, and TOB1 sites, mutant PUM constructs were generated where TGT was mutated to ACA in the PUM motif. For microRNA inhibitor luciferase reporter experiments with RRAGD and TOB1 sites, mutant miRNA seed constructs were generated where three nucleotides in position 2-7 were mutated. In all cases, constructs were assembled as described above.

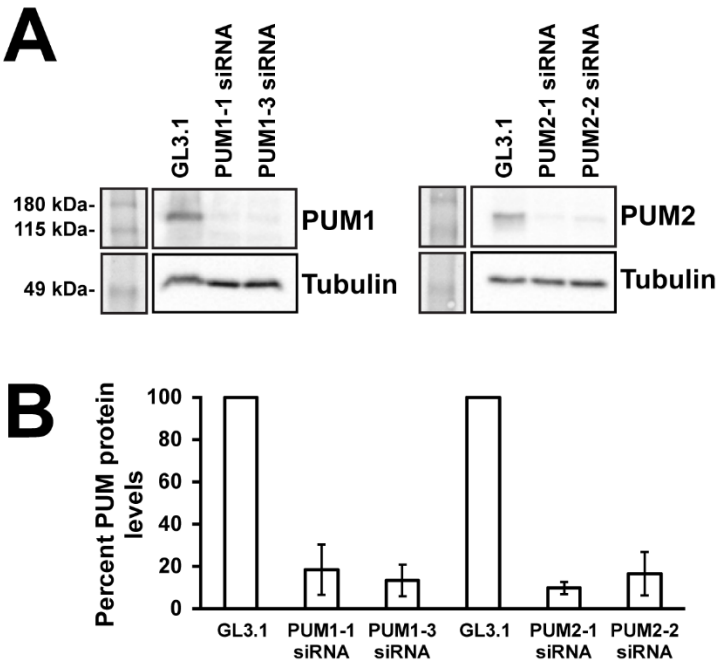
**Plasmid transfection and luciferase assays**

T-REx-293 and PUM DKO cells were seeded in 96-well plates and transfected in technical triplicate at 70% confluency. For all experiments, WT and mutant plasmids were transfected in parallel. For transfections of the initial 26 candidate set, TransIT-LT1 reagent (Mirus) was used per manufacturer's instruction to add 10 ng of reporter plasmid to cells. In miRNA inhibitor experiments, calcium phosphate transfection was used to transfect 10 ng of reporter plasmid with 0.75 μM of each inhibitor. Anti-miR-30, anti-miR-25, and the control hairpin inhibitors were manufactured by Dharmacon. Anti-miR-17 family LNA

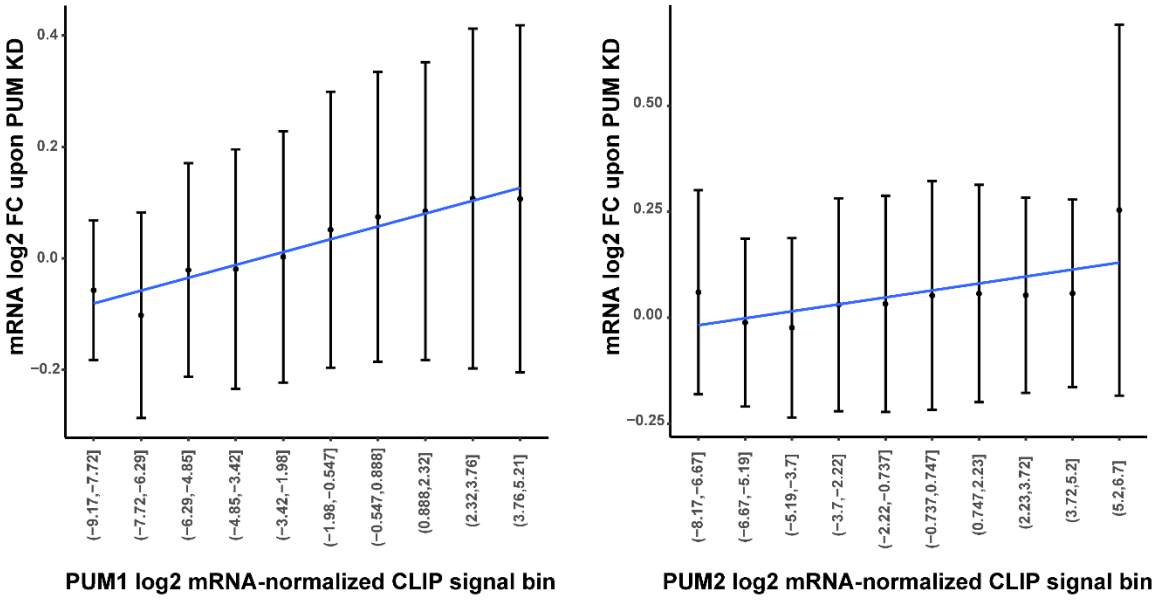
and its control were obtained from Exiqon. Cells were lysed 24 hours after transfection with 20 µl of Passive Lysis Buffer (Promega). Five microliters of cell lysate were used for dual luciferase reporter measurements (Promega). Luciferase substrates were diluted 1:5 in use. Renilla luciferase signal was normalized to the firefly luciferase signal produced from the same plasmid to control for transfection efficiency. For each experiment, at least three biological replicates were performed. Experiments with greater than 50% coefficients of variability between the technical replicates were omitted from downstream analysis. The ratio of normalized WT and mutant luciferase was calculated to determine the effect of site mutation on gene expression. Comparisons between WT and mutant constructs were analyzed by two-tailed paired t-test. Comparisons of WT/mutant ratios between cell conditions were analyzed by two-tailed Welch's t-test. p-value significance was defined at 0.05.

The sequencing raw data and processed binding site data is available at the NCBI Gene Expression Omnibus with accession GSE110520.

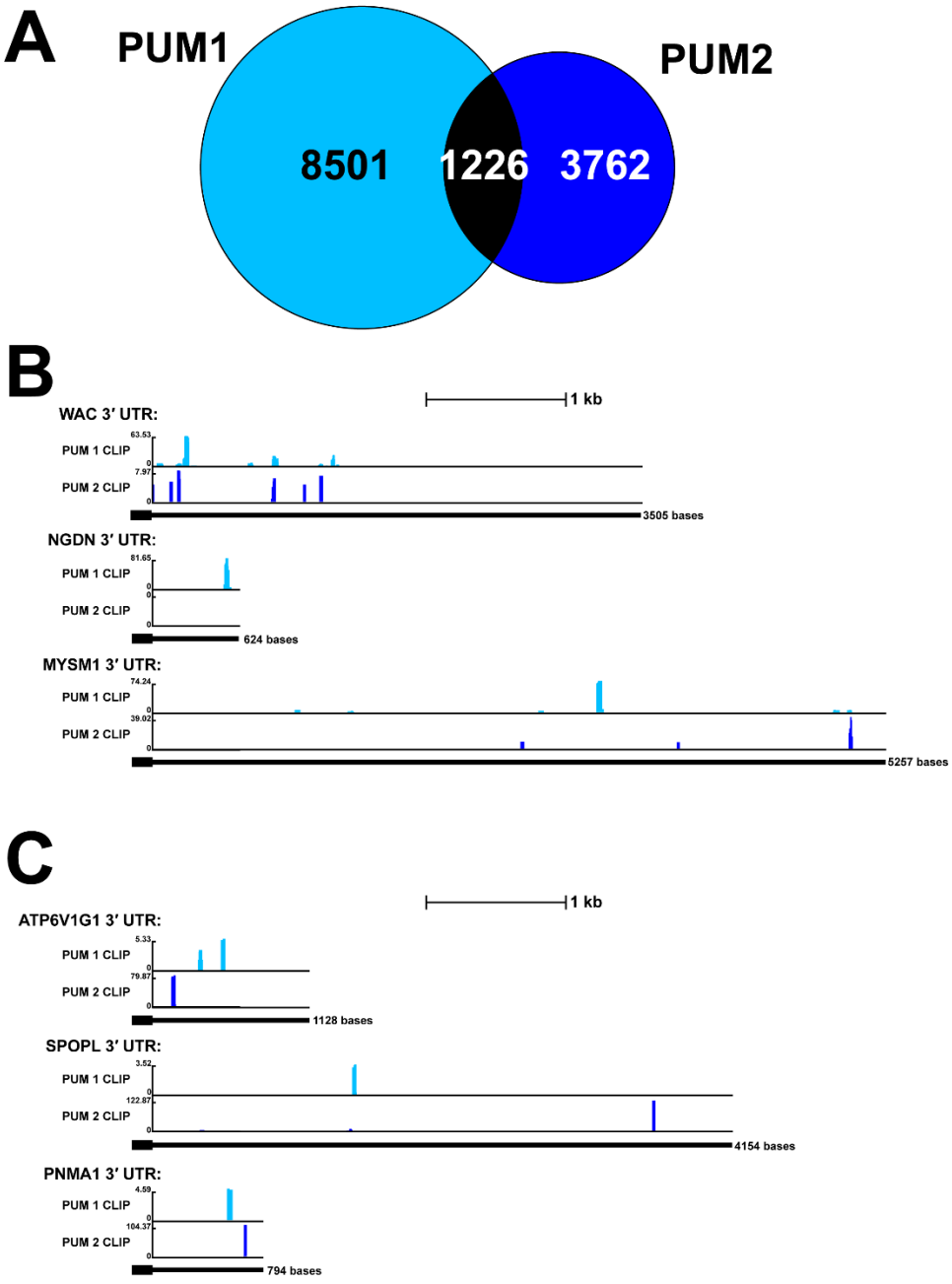
## Supplementary Figures and Tables



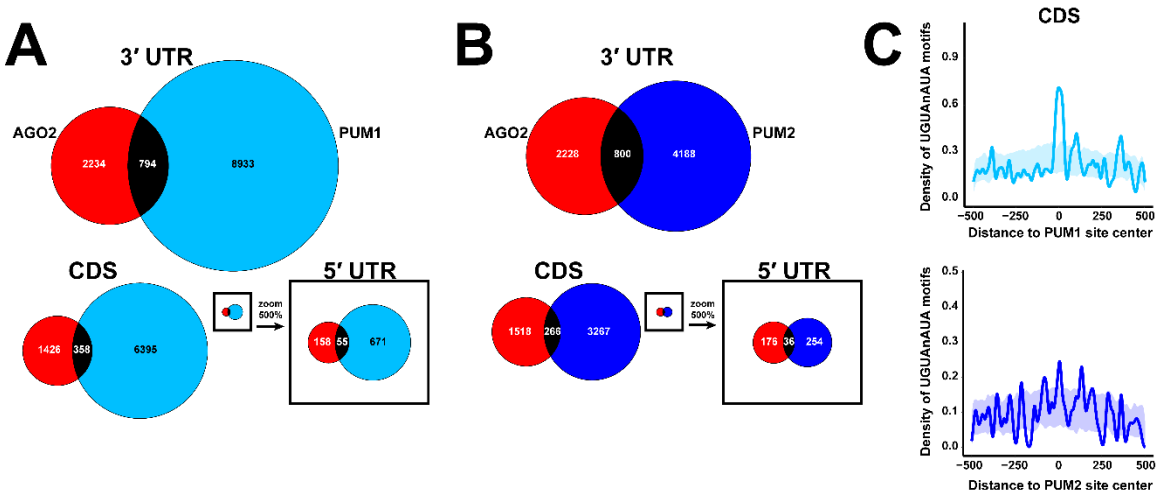

**Supplemental Figure 2.1: PUM1 and PUM2 siRNA knockdown for AGO2-CLIP. A)** Western blot of PUM1 and PUM2 knockdown with two separate siRNAs each. An siRNA targeting firefly luciferase (GL3.1) was used as a control. **B)** Percent PUM protein after knockdown of PUM1 (n=2) and PUM2 (n=4). Each condition was normalized to total tubulin and then compared to the GL3.1 control.



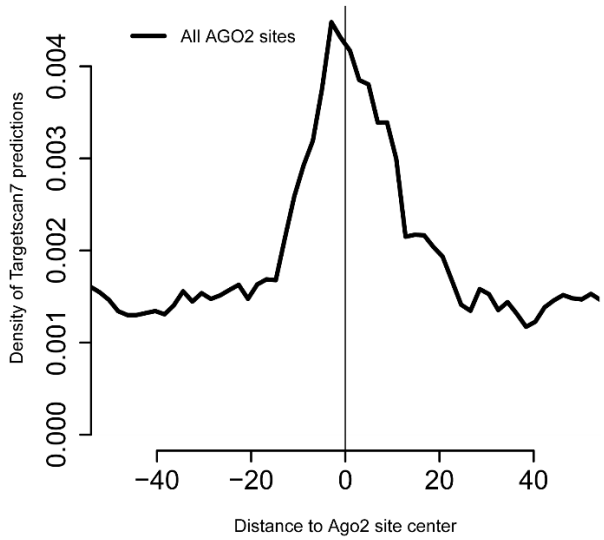

**Supplemental Figure 2.2. Quantitative relationship between PUM CLIP signal strength and effects on mRNA levels.** The sum of PUM1 or PUM2 CLIP readcounts for each 3' UTR, corrected for library depth by DESeq2, normalized to mRNA levels [85], and log2 transformed, were binned. For the transcripts in each bin, the mean mRNA $\log_2$ fold change upon PUM KD [52] was calculated. A) Correlation between mRNA level changes upon PUM KD and PUM1 CLIP signal. Pearson correlation coefficient r = 0.959, p = 1.23 x $10^{-5}$. B) Same, for PUM2 CLIP signal; r = 0.658, p = 0.0384.



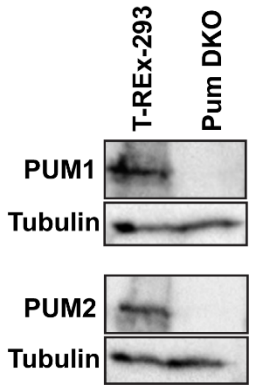

**Supplemental Figure 2.3: PUM1 and PUM2 display differential binding. A)** Number of sites within 3′ UTRs occupied by PUM1 (light blue), PUM2 (dark blue), or both (black). **B)** Example 3' UTRs containing sites with strong binding signal for PUM1 (light blue) but not PUM2 (dark blue). Peaks of interest are depicted with a black arrow. **C)** Example 3' UTRs containing sites with strong binding signal for PUM2 (dark blue) but not PUM1 (light blue). Peaks of interest are depicted with a black arrow.



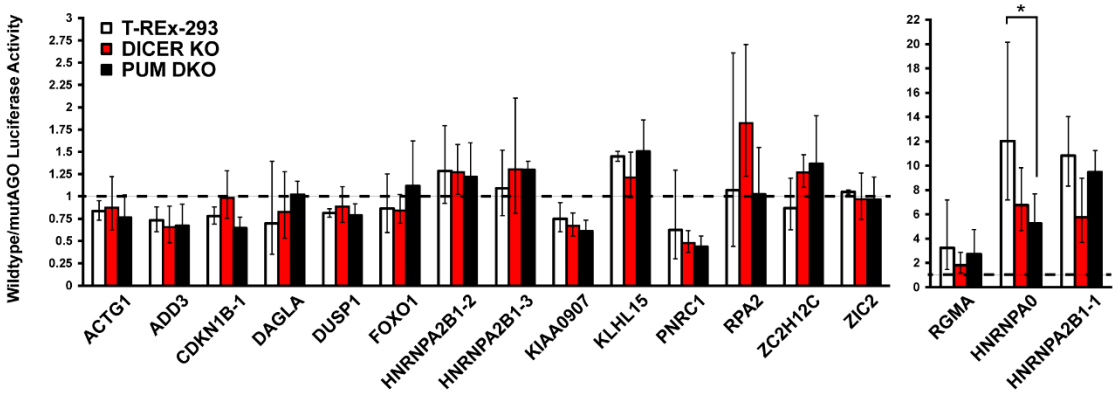

**Supplemental Figure 2.4: Venn Diagram of overlapping PUM and AGO2 sites. A)** Number of sites within 3′ UTRs, the CDS, and 5'UTRs occupied by AGO2 (red), PUM1 (light blue), or both (black). **B)** Number of sites within 3′ UTRs, the CDS, and 5'UTRs occupied by AGO2 (red), PUM2 (dark blue), or both (black). **C)** Density of PUM motifs surrounding PUM1 (light blue) and PUM2 (dark blue) sites. The shaded areas represent randomized controls for the corresponding populations.



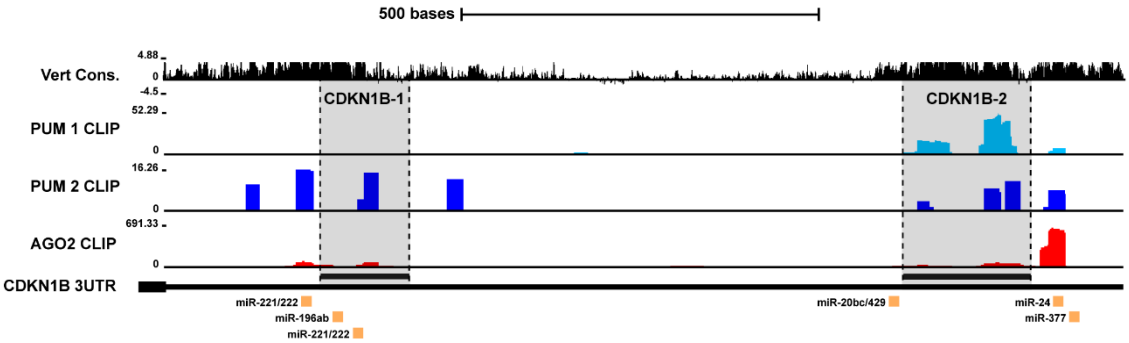

**Supplemental Figure 2.5: AGO2 peak centers overlap with predicted miRNA seed sites.** Density of Targetscan7 predictions surrounding all AGO2 peaks. A vertical black line denoted the AGO2 peak center.



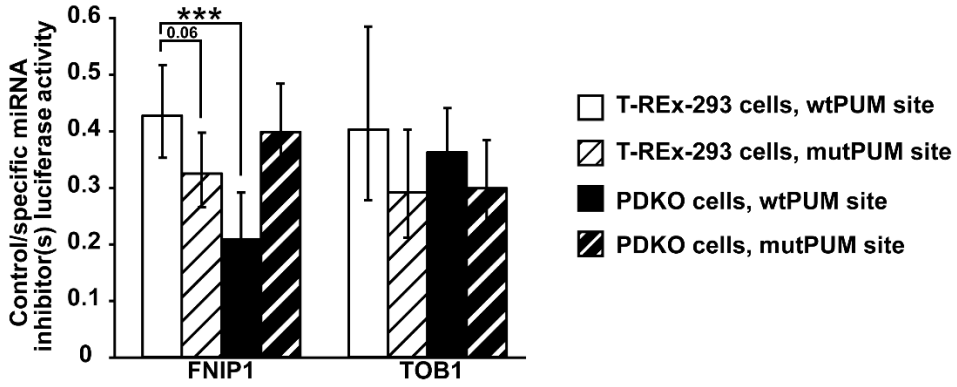

**Supplemental Figure 2.6: Western blot of PUM1 and PUM2 in T-REx-293 and PUM double knockout cells.** Tubulin was used as a loading control.



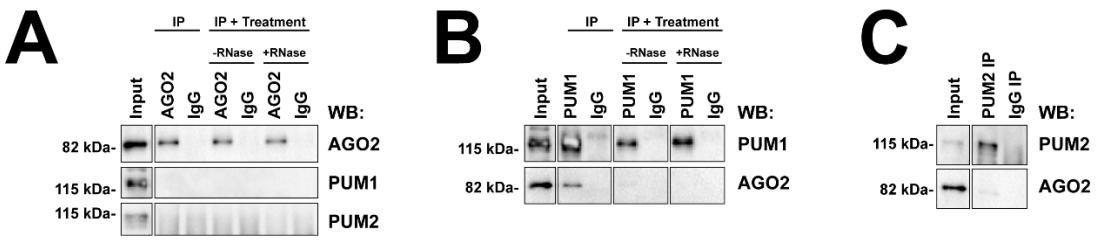

**Supplemental Figure 2.7: Candidate sites that were not examined further.** A subset of candidates showed no regulation in all cell conditions (ACTG1, DAGLA, FOXO1, HNRNPA2B-2, HNRNPA2B1-3, ZC3H12C, ZIC2), regulation that was not significantly dependent on AGO or PUM (ADD3, CDKN1B, DUSP1, KIA0907, KLHL15, PNRC1, RGMA, HNRNPA2B1-1), or excessively high protective site activity (HNRNPA0). *= $p<0.05$.



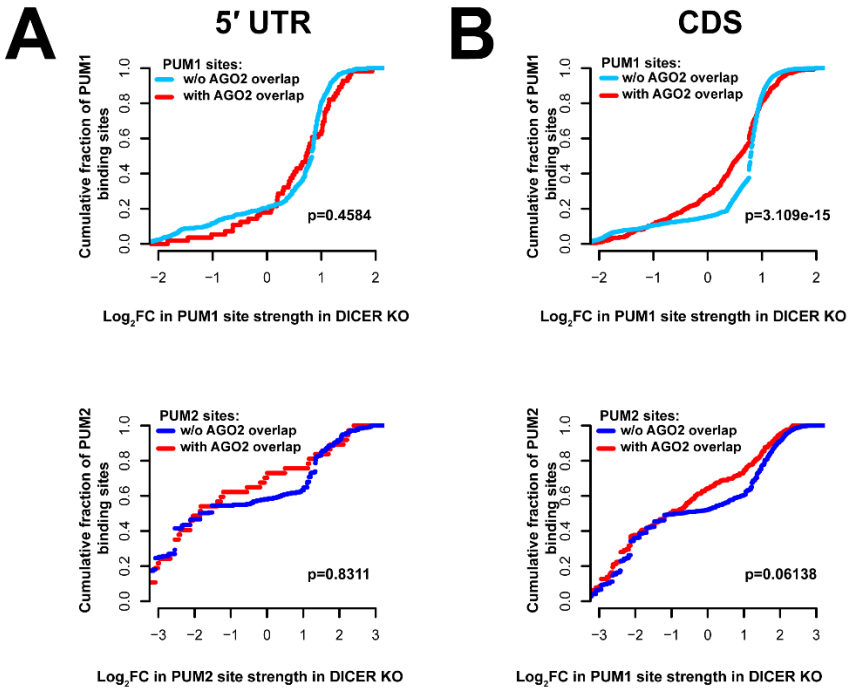

**Supplemental Figure 2.8: UCSC browser view displaying the CDKN1B 3' UTR.** Tracks show the Vertebrate conservation score (black) as well as CLIP data for PUM1 (light blue), PUM2 (dark blue), and AGO2 (red). Candidate regions CDKN1B-1 and CDKN1B-2 are shaded gray.




**Supplemental Figure 2.9: Two sites show incomplete or no PUM dependent effects on miRNA dependent regulation.** Luciferase reporter constructs co-transfected with control or specific miRNA inhibitors were tested under wildtype (white) and PUM double knockout (black) conditions with WT (solid) or mutant PUM sites (stripes). *= $p<0.05$, **= $p<0.01$, ***= $p<0.001$.




**Supplemental Figure 2.10: Immunoprecipitation of AGO2, pum1, and PUM2 shows no evidence of strong direct interactions. A)** Immunoblot of AGO2 immunoprecipitation (IP) using antibodies against AGO2, PUM1, and PUM2. IP complexes were treated with PBS (-RNase) or with 50µg/mL RNase A (+RNase). Species matched nonspecific IgG was used as a negative control. **B)** Immunoblot of PUM1 immunoprecipitation (IP) using antibodies against PUM1 and AGO2. IP complexes were treated with PBS (-RNase) or with 50ug/mL RNase A (+RNase). Species matched nonspecific IgG was used as a negative control. **C)** Immunoblot of PUM2 immunoprecipitation using antibodies against PUM2 and AGO2. Species matched nonspecific IgG was used as a negative control.




**Supplemental Figure 2.11: AGO2-PUM site co-occupancy affects PUM1 binding in CDS sites, but not PUM2 binding in the CDS or PUM1 and PUM2 binding in the 5'UTR. A)** Cumulative distribution plots of the log fold change in PUM site strength, with AGO2 overlap (red) and without AGO2 overlap (PUM1: light blue, PUM2: dark blue) within 5'UTRs. **B)** Cumulative distribution plots of the log fold change in PUM site strength, with AGO2 overlap (red) and without AGO2 overlap (PUM1: light blue, PUM2: dark blue) within the CDS.

## References


1 Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233, doi:10.1016/j.cell.2009.01.002 (2009).

2 Meister, G. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* **14**, 447-459, doi:10.1038/nrg3462 (2013).

3 Iwakawa, H. O. & Tomari, Y. The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol.* **25**, 651-665, doi:10.1016/j.tcb.2015.07.011 (2015).

4 Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16**, 421-433, doi:10.1038/nrg3965 (2015).

5 Djuranovic, S., Nahvi, A. & Green, R. A parsimonious model for gene regulation by miRNAs. *Science* **331**, 550-553, doi:10.1126/science.1191138 (2011).

6 Fabian, M. R., Sonenberg, N. & Filipowicz, W. Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* **79**, 351-379, doi:10.1146/annurev-biochem-060308-103103 (2010).

7 Friedman, R. C., Farh, K. K., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92-105, doi:10.1101/gr.082701.108 (2009).

8 Wickens, M., Bernstein, D. S., Kimble, J. & Parker, R. A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet.* **18**, 150-157 (2002).

9 Zhang, B. *et al.* A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* **390**, 477-484, doi:10.1038/37297 (1997).

10 Zamore, P. D., Williamson, J. R. & Lehmann, R. The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**, 1421-1433 (1997).

11 Wang, X., McLachlan, J., Zamore, P. D. & Hall, T. M. Modular recognition of RNA by a human pumilio-homology domain. *Cell* **110**, 501-512 (2002).

12 Wang, X., Zamore, P. D. & Hall, T. M. Crystal structure of a Pumilio homology domain. *Mol. Cell* **7**, 855-865 (2001).

13 Morris, A. R., Mukherjee, N. & Keene, J. D. Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets. *Mol. Cell. Biol.* **28**, 4093-4103, doi:10.1128/MCB.00155-08 (2008).

14 Hall, T. M. De-coding and re-coding RNA recognition by PUF and PPR repeat proteins. *Curr. Opin. Struct. Biol.* **36**, 116-121, doi:10.1016/j.sbi.2016.01.010 (2016).

15 Lehmann, R. & Nüsslein-Volhard, C. Involvement of the pumilio gene in the transport of an abdominal signal in the Drosophila embryo. *Nature* **329**, 167, doi:10.1038/329167a0 (1987).

16 Tautz, D. Regulation of the Drosophila segmentation gene hunchback by two maternal morphogenetic centres. *Nature* **332**, 281-284, doi:10.1038/332281a0 (1988).
17 Barker, D. D., Wang, C., Moore, J., Dickinson, L. K. & Lehmann, R. Pumilio is essential for function but not for distribution of the Drosophila abdominal determinant Nanos. *Genes Dev.* **6**, 2312-2326 (1992).
18 Murata, Y. & Wharton, R. P. Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in Drosophila embryos. *Cell* **80**, 747-756 (1995).
19 Dubnau, J. *et al.* The staufen/pumilio pathway is involved in Drosophila long-term memory. *Curr. Biol.* **13**, 286-296 (2003).
20 Menon, K. P. *et al.* The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E. *Neuron* **44**, 663-676, doi:10.1016/j.neuron.2004.10.028 (2004).
21 Ye, B. *et al.* Nanos and Pumilio are essential for dendrite morphogenesis in Drosophila peripheral neurons. *Curr. Biol.* **14**, 314-321, doi:10.1016/j.cub.2004.01.052 (2004).
22 Lin, H. & Spradling, A. C. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development* **124**, 2463-2476 (1997).
23 Parisi, M. & Lin, H. The Drosophila pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis. *Genetics* **153**, 235-250 (1999).
24 Crittenden, S. L. *et al.* A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans. *Nature* **417**, 660-663, doi:10.1038/nature754 (2002).
25 Forbes, A. & Lehmann, R. Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. *Development* **125**, 679-690 (1998).
26 Galgano, A. *et al.* Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. *PLoS One* **3**, e3164, doi:10.1371/journal.pone.0003164 (2008).
27 Zhang, M. *et al.* Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. *Genes Dev.*, doi:10.1101/gad.298752.117 (2017).
28 Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129-141, doi:10.1016/j.cell.2010.03.009 (2010).
29 Spassov, D. S. & Jurecic, R. Cloning and comparative sequence analysis of PUM1 and PUM2 genes, human members of the Pumilio family of RNA-binding proteins. *Gene* **299**, 195-204 (2002).
30 Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419, doi:10.1126/science.1260419 (2015).
31 Chen, D. *et al.* Pumilio 1 suppresses multiple activators of p53 to safeguard spermatogenesis. *Curr. Biol.* **22**, 420-425, doi:10.1016/j.cub.2012.01.039 (2012).

32 Xu, E. Y., Chang, R., Salmon, N. A. & Reijo Pera, R. A. A gene trap mutation of a murine homolog of the Drosophila stem cell factor Pumilio results in smaller testes but does not affect litter size or fertility. *Mol. Reprod. Dev.* **74**, 912-921, doi:10.1002/mrd.20687 (2007).
33 Mak, W., Fang, C., Holden, T., Dratver, M. B. & Lin, H. An Important Role of Pumilio 1 in Regulating the Development of the Mammalian Female Germline. *Biol. Reprod.* **94**, 134, doi:10.1095/biolreprod.115.137497 (2016).
34 Vessey, J. P. *et al.* Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3222-3227, doi:10.1073/pnas.0907128107 (2010).
35 Driscoll, H. E., Muraro, N. I., He, M. & Baines, R. A. Pumilio-2 regulates translation of Nav1.6 to mediate homeostasis of membrane excitability. *J. Neurosci.* **33**, 9644-9654, doi:10.1523/JNEUROSCI.0921-13.2013 (2013).
36 Siemen, H., Colas, D., Heller, H. C., Brustle, O. & Pera, R. A. Pumilio-2 function in the mouse nervous system. *PLoS One* **6**, e25932, doi:10.1371/journal.pone.0025932 (2011).
37 Gennarino, V. A. *et al.* Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels. *Cell* **160**, 1087-1098, doi:10.1016/j.cell.2015.02.012 (2015).
38 Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **164**, 69-80, doi:10.1016/j.cell.2015.12.017 (2016).
39 Kedde, M. *et al.* A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol.* **12**, 1014-1020, doi:10.1038/ncb2105 (2010).
40 Miles, W. O., Tschop, K., Herr, A., Ji, J. Y. & Dyson, N. J. Pumilio facilitates miRNA regulation of the E2F3 oncogene. *Genes Dev.* **26**, 356-368, doi:10.1101/gad.182568.111 (2012).
41 Van Etten, J. *et al.* Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *J. Biol. Chem.* **287**, 36370-36383, doi:10.1074/jbc.M112.373522 (2012).
42 Goldstrohm, A. C., Hook, B. A., Seay, D. J. & Wickens, M. PUF proteins bind Pop2p to regulate messenger RNAs. *Nat. Struct. Mol. Biol.* **13**, 533-539, doi:10.1038/nsmb1100 (2006).
43 Weidmann, C. A., Raynard, N. A., Blewett, N. H., Van Etten, J. & Goldstrohm, A. C. The RNA binding domain of Pumilio antagonizes poly-adenosine binding protein and accelerates deadenylation. *RNA* **20**, 1298-1319, doi:10.1261/rna.046029.114 (2014).
44 Bohn, J. A. *et al.* Identification of diverse target RNAs that are functionally regulated by human Pumilio proteins. *Nucleic Acids Res.* **46**, 362-386, doi:10.1093/nar/gkx1120 (2018).
45 Chritton, J. J. & Wickens, M. Translational repression by PUF proteins in vitro. *RNA* **16**, 1217-1225, doi:10.1261/rna.2070110 (2010).

46 Cao, Q., Padmanabhan, K. & Richter, J. D. Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition. *RNA* **16**, 221-227, doi:10.1261/rna.1884610 (2010).
47 Cho, P. F. *et al.* Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. *Curr. Biol.* **16**, 2035-2041, doi:10.1016/j.cub.2006.08.093 (2006).
48 Archer, S. K., Luu, V. D., de Queiroz, R. A., Brems, S. & Clayton, C. Trypanosoma brucei PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. *PLoS Pathog.* **5**, e1000565, doi:10.1371/journal.ppat.1000565 (2009).
49 Kaye, J. A., Rose, N. C., Goldsworthy, B., Goga, A. & L'Etoile, N. D. A 3'UTR pumilio-binding element directs translational activation in olfactory sensory neurons. *Neuron* **61**, 57-70, doi:10.1016/j.neuron.2008.11.012 (2009).
50 Suh, N. *et al.* FBF and its dual control of gld-1 expression in the Caenorhabditis elegans germline. *Genetics* **181**, 1249-1260, doi:10.1534/genetics.108.099440 (2009).
51 Pique, M., Lopez, J. M., Foissac, S., Guigo, R. & Mendez, R. A combinatorial code for CPE-mediated translational control. *Cell* **132**, 434-448, doi:10.1016/j.cell.2007.12.038 (2008).
52 HafezQorani, S. *et al.* Modeling the combined effect of RNA-binding proteins and microRNAs in post-transcriptional regulation. *Nucleic Acids Res.* **44**, e83, doi:10.1093/nar/gkw048 (2016).
53 Jiang, P. & Coller, H. Functional interactions between microRNAs and RNA binding proteins. *Microrna* **1**, 70-79 (2012).
54 Jiang, P., Singh, M. & Coller, H. A. Computational assessment of the cooperativity between RNA binding proteins and MicroRNAs in Transcript Decay. *PLoS Comput. Biol.* **9**, e1003075, doi:10.1371/journal.pcbi.1003075 (2013).
55 Leibovich, L., Mandel-Gutfreund, Y. & Yakhini, Z. A structural-based statistical approach suggests a cooperative activity of PUM1 and miR-410 in human 3'-untranslated regions. *Silence* **1**, 17, doi:10.1186/1758-907X-1-17 (2010).
56 Dueck, A., Ziegler, C., Eichner, A., Berezikov, E. & Meister, G. microRNAs associated with the different human Argonaute proteins. *Nucleic Acids Res.* **40**, 9850-9862, doi:10.1093/nar/gks705 (2012).
57 Wang, D. *et al.* Quantitative functions of Argonaute proteins in mammalian development. *Genes Dev.* **26**, 693-704, doi:10.1101/gad.182758.111 (2012).
58 Landthaler, M. *et al.* Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* **14**, 2580-2596, doi:10.1261/rna.1351608 (2008).
59 Li, Y., Estep, J. A. & Karginov, F. V. Transcriptome-wide Identification and Validation of Interactions between the miRNA Machinery and HuR on mRNA Targets. *J. Mol. Biol.* **430**, 285-296, doi:10.1016/j.jmb.2017.12.006 (2018).

60 Bogerd, H. P., Whisnant, A. W., Kennedy, E. M., Flores, O. & Cullen, B. R. Derivation and characterization of Dicer- and microRNA-deficient human cells. *RNA* **20**, 923-937, doi:10.1261/rna.044545.114 (2014).
61 Kishore, S. *et al.* A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat Methods* **8**, 559-564, doi:10.1038/nmeth.1608 (2011).
62 Memczak, S. *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333-338, doi:10.1038/nature11928 (2013).
63 Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **153**, 654-665, doi:10.1016/j.cell.2013.03.043 (2013).
64 Li, J. H., Liu, S., Zhou, H., Qu, L. H. & Yang, J. H. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **42**, D92-97, doi:10.1093/nar/gkt1248 (2014).
65 Bailey, T. L. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653-1659, doi:10.1093/bioinformatics/btr261 (2011).
66 Konig, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17**, 909-915, doi:10.1038/nsmb.1838 (2010).
67 Loeb, G. B. *et al.* Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. *Mol. Cell* **48**, 760-770, doi:10.1016/j.molcel.2012.10.002 (2012).
68 Plass, M., Rasmussen, S. H. & Krogh, A. Highly accessible AU-rich regions in 3' untranslated regions are hotspots for binding of regulatory factors. *PLoS Comput. Biol.* **13**, e1005460, doi:10.1371/journal.pcbi.1005460 (2017).
69 Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, doi:10.7554/eLife.05005 (2015).
70 Sternburg, E. L., Dias, K. C. & Karginov, F. V. Selection-dependent and Independent Generation of CRISPR/Cas9-mediated Gene Knockouts in Mammalian Cells. *J Vis Exp*, doi:10.3791/55903 (2017).
71 Frohn, A. *et al.* Dicer-dependent and -independent Argonaute2 protein interaction networks in mammalian cells. *Mol. Cell. Proteomics* **11**, 1442-1456, doi:10.1074/mcp.M112.017756 (2012).
72 Naudin, C. *et al.* PUMILIO/FOXP1 signaling drives expansion of hematopoietic stem/progenitor and leukemia cells. *Blood* **129**, 2493-2506, doi:10.1182/blood-2016-10-747436 (2017).
73 Lee, C. D. & Tu, B. P. Glucose-Regulated Phosphorylation of the PUF Protein Puf3 Regulates the Translational Fate of Its Bound mRNAs and Association with RNA Granules. *Cell Rep.* **11**, 1638-1650, doi:10.1016/j.celrep.2015.05.014 (2015).
74 Boycott, K. M. *et al.* Homozygous deletion of the very low density lipoprotein receptor gene causes autosomal recessive cerebellar hypoplasia with cerebral

gyral simplification. *Am. J. Hum. Genet.* **77**, 477-483, doi:10.1086/444400 (2005).
75 Trommsdorff, M. *et al.* Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. *Cell* **97**, 689-701 (1999).
76 Schlotawa, L. *et al.* Cerebellar ataxia, mental retardation and dysequilibrium syndrome 1 (CAMRQ1) caused by an unusual constellation of VLDLR mutation. *J. Neurol.* **260**, 1678-1680, doi:10.1007/s00415-013-6941-z (2013).
77 Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I. & Filipowicz, W. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125**, 1111-1124, doi:10.1016/j.cell.2006.04.031 (2006).
78 Young, L. E., Moore, A. E., Sokol, L., Meisner-Kober, N. & Dixon, D. A. The mRNA stability factor HuR inhibits microRNA-16 targeting of COX-2. *Mol. Cancer Res.* **10**, 167-180, doi:10.1158/1541-7786.MCR-11-0337 (2012).
79 Bottini, S. *et al.* Post-transcriptional gene silencing mediated by microRNAs is controlled by nucleoplasmic Sfpq. *Nat Commun* **8**, 1189, doi:10.1038/s41467-017-01126-x (2017).
80 Xue, Y. *et al.* Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell* **152**, 82-96, doi:10.1016/j.cell.2012.11.045 (2013).
81 Kedde, M. *et al.* RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell* **131**, 1273-1286, doi:10.1016/j.cell.2007.11.034 (2007).
82 Iadevaia, V. & Gerber, A. P. Combinatorial Control of mRNA Fates by RNA-Binding Proteins and Non-Coding RNAs. *Biomolecules* **5**, 2207-2222, doi:10.3390/biom5042207 (2015).
83 Lebedeva, S. *et al.* Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol. Cell* **43**, 340-352, doi:10.1016/j.molcel.2011.06.008 (2011).
84 Mukherjee, N. *et al.* Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol. Cell* **43**, 327-339, doi:10.1016/j.molcel.2011.06.007 (2011).
85 Karginov, F. V. & Hannon, G. J. Remodeling of Ago2-mRNA interactions upon cellular stress reflects miRNA complementarity and correlates with altered translation rates. *Genes Dev.* **27**, 1624-1632, doi:10.1101/gad.215939.113 (2013).

# CHAPTER 3

## The Role of Mammalian Pumilio Proteins in Regulating Cell Adhesion and Migration Pathways


## Abstract

Pumilio proteins belong to a broad group of repressive RNA-binding proteins that control mRNA translation and stability by binding to the 3' UTR of target mRNAs. Mammals have two canonical Pumilio proteins, PUM1 and PUM2, which are known to act in many biological processes, including embryonic development, neurogenesis, cell cycle regulation and genomic stability. Here, we characterized a new role of both PUM1 and PUM2 in regulating cell adhesion and migration. PUM double knockout (DKO) T-REx-293 cells had defects in growth rate and cell morphology. PUM DKO cells grew in clumps, which arose from an inability to escape cell-cell contacts. PUM DKO cells had a collective cell migration rate significantly lower than that of WT cells, and displayed changes in actin morphology. Finally, gene ontology analysis of differentially expressed genes in PUM DKO cells for both cellular component and biological process showed enrichment in categories related to adhesion and migration. This study expands our current understanding of mammalian PUM proteins, which can aid in developing better models for its function in both developmental processes and disease.

## Introduction

Post-transcriptional regulation is fundamental to proper control of protein expression. This control is executed by trans-acting RNA binding factors, including RNA binding proteins (RBPs) and small RNAs. Pumilio proteins belong to a broad group of RBPs that control mRNA translation and stability by binding to the 3' UTR. Members of the highly conserved PUF family of RNA proteins, Pumilio proteins have been studied in organisms from yeast to humans[1]. Pumilio was originally characterized in Drosophila as a key regulator in embryonic development[2-4], where it acts cooperatively with nanos in order to regulate protein expression of the embryonic patterning gene, hunchback. Since its discovery, Pumilio proteins have been extensively characterized in invertebrates, including flies, worms, and yeast. These studies have highlighted the evolutionarily conserved function of Pumilio proteins in regulating development and germline maintenance[5-8].

Mammals have two canonical Pumilio proteins, PUM1 and PUM2, which are highly conserved across organisms. The two mammalian homologs share 69% identity / 74% similarity along their entire length. At the C-terminus is the ~340 amino acid Pumilio homology domain (Pum-HD) that is conserved among Pumilio homologs across organisms. Much of the differences in Pumilio proteins in different organisms, as well as between the two human paralogs, lies within the N-terminal region. PUM1 contains an extended N-terminal region, which has been hypothesized to recruit other binding partners and be the site of additional regulation. Like in many RBPs, significant portions of the N-terminal regions of both PUMs are predicted to be disordered, which likely play a role in recruitment to phase-separated droplets under various cellular conditions[9]. The two

proteins are largely redundant, although differences in target binding do exist[10-12]. Both proteins recognize target transcripts by binding to a conserved Pumilio recognition element (PRE) 5′-UGUAnAUA-3′[13-15]. When bound, Pumilio proteins repress target transcripts through the recruitment of machinery which can inhibit translation and destabilize the transcript[11,16]. Similar to Drosophila, NANOS paralogues are thought to play a role in modulating the binding and regulatory behavior of PUMs[17,18].

In recent years, the role of Pumilio proteins in mammals has been studied, which has validated many conserved functions for Pum proteins as well as identified new ones. PUM proteins are essential to mammalian embryonic development; double knockout mice embryos are unable to complete gastrulation and are embryonic lethal[19,20]. PUM is also present in mouse oocytes and is demonstrated to have a maternal effect during early embryogenesis[21]. Pumilio proteins also play an important role in gametogenesis[22,23]. For both oogenesis and spermatogenesis, Pumilio proteins regulate transcripts that control self-renewal of early germ cell populations.

Pumilio proteins play a substantial role in neurogenesis and in proper neuron function. Brain specific PUM double knockout mice show defects in brain development[24] and PUM2 has been demonstrated to help specify neuronal cell fate in neuronal stem cells[25]. Loss of PUM1 has been shown to increase the expression of Ataxin1, a defect that in both mice and humans causes spinocerebellar ataxia[26,27]. Within the immune system, Pumilio proteins are involved in the regulation of genes involved in innate immunity[28] and in the proper maintenance of hematopoietic stem cells[29]. In addition, Pum proteins are known to

regulate transcripts of proteins necessary for cell cycle regulation and genome stability[30-33].

A role of PUM proteins in adhesion and migration has not yet been described, although evidence exists from CLIP, RIP, and RNA-seq datasets. A recent paper which analyzed multiple datasets in order to identify PUM targets found an enrichment of cell adhesion and migration related transcripts under regulation of PUM proteins[11]. This was confirmed further in a recent review which looked at PUM target transcripts in both cultured cells and neuronal tissues[34]. Here, we propose a role of mammalian PUM proteins in regulating cell adhesion and migration though characterization of PUM double knockout (DKO) T-REx-293 cells, generated as described in a previously published study[10]. We validate that the observed phenotype is PUM dependent and provide further evidence that PUM proteins regulate transcripts important in cell adhesion and migration pathways.

## Results

### Double knockout of PUM1 and PUM2 affects the growth rate of T-REx-293 cells

In agreement with previous PUM knockout experiments[32], PUM DKO cells displayed a diminished growth rate compared to WT (Figure 3.1A) (Table 3.1). This growth phenotype is also observed in independently generated HCT116 PUM DKO cells[32] (Supplementary Figure 3.1A).To confirm that the effect on growth rate is PUM dependent, PUM1 and PUM2 were independently reintroduced into PUM DKO cells, stable integrant cell populations were selected, and PUM expression was confirmed by western blot (Figure 3.1B). PUM1 rescue increased growth to a level that is significantly higher than that of PUM DKO (Figure 3.1C) (Table 3.1). PUM2 showed a similar increase, although this increase is not statistically significant. However, rescue of either PUM1 or PUM2 independently did not fully restore the growth rate to WT levels. Dosage effects of total PUM protein have been reported to affect neurodevelopment and genomic stability[26,27,35], and could explain why neither single rescue can fully restore WT growth rates. This observation is also in agreement with the notion that PUM proteins have non-redundant functions, with both proteins being necessary for WT function to be restored.

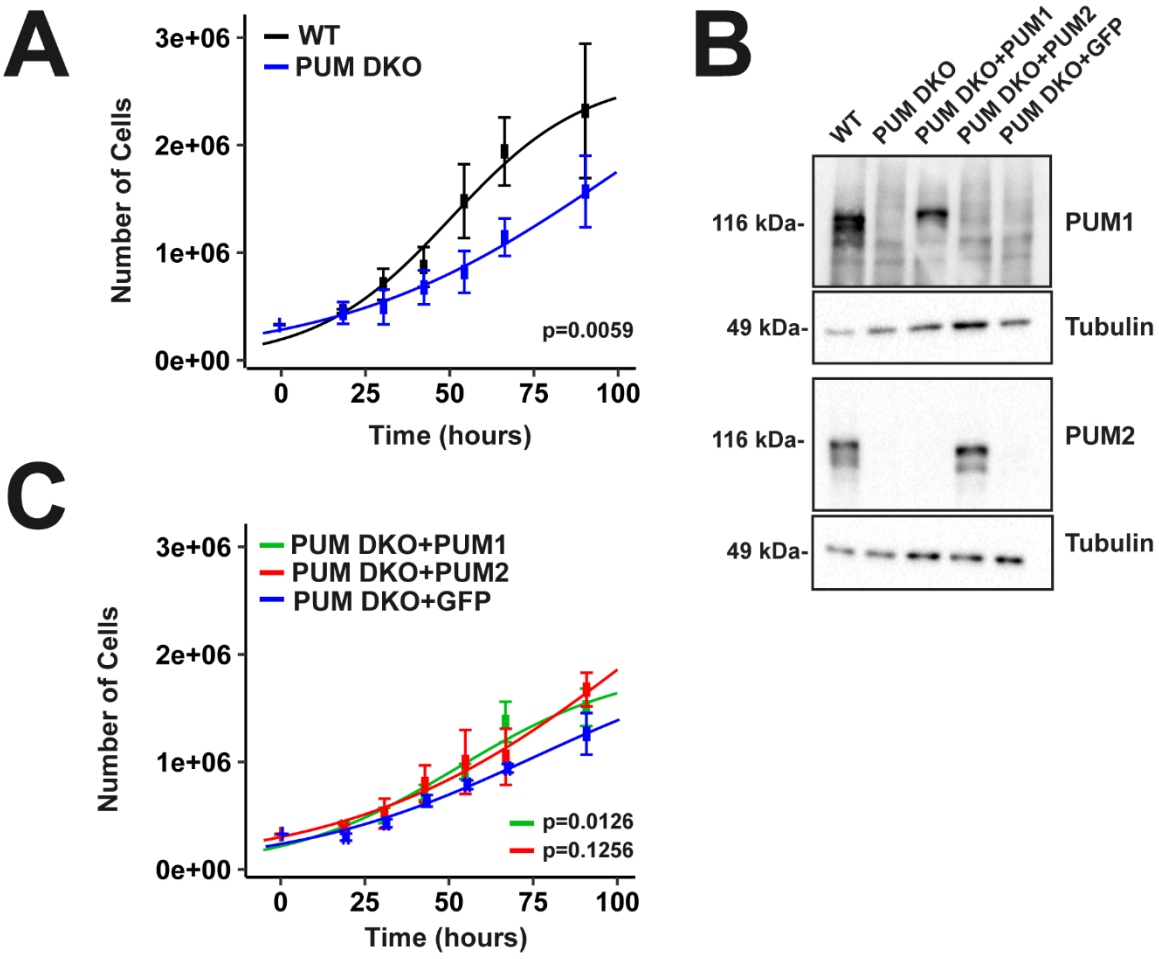



**Figure 3.1: Cell growth rate is modulated by both PUM proteins.** (A) Growth rates of WT (black) and PUM DKO (blue) T-REx-293 cells. (B) Western blot for PUM1 and PUM2 in WT, PUM DKO, and T-REx-293 cells stably transfected with PUM1, PUM2, or GFP. Tubulin serves as a control. (C) Growth rates of PUM DKO T-REx-293 cells stably transfected with PUM1 (green), PUM2 (red), or GFP (blue). Growth measurements were calculated from three biological replicates, with error bars representing standard error of the mean. For the purposes of plotting, the time points are grouped for each cell type / replicate group (since the data was collected at slightly different timepoints for each replicate), generating error bars along the x axis direction (standard error of the mean). Logistic growth model fits of the data were compared using ANOVA F-test to determine statistical significance.

**Table 3.1: Growth rate equations for T-Rex-293 and HCT116 cells.**

| | Nmax +/- SE (# cells) | Tmid +/- SE (hours) | Max doubling time +/- SE (hours) |
|---|---|---|---|
| T-Rex-293 WT | 2600000 +/- 680000 | 51 +/- 14 | 14 +/- 5.9 |
| T-Rex-293 PUM DKO | 3500000 +/- 5500000 | 99 +/- 120 | 28 +/- 19 |
| T-Rex-293 PUM DKO + GFP | 2100000 +/- 850000 | 75 +/- 30 | 25 +/- 6.7 |
| T-Rex-293 PUM DKO + PUM1 | 1900000 +/- 430000 | 53 +/- 14 | 18 +/- 5.3 |
| T-Rex-293 PUM DKO + PUM2 | 4100000 +/- 7600000 | 108 +/- 139 | 30 +/- 19 |
| HCT116 WT | 3000000 +/- 290000 | 42 +/- 4.3 | 8.9 +/- 2.3 |
| HCT116 PUM DKO | 2400000 +/- 360000 | 46 +/- 7.6 | 12 +/- 3.7 |

## PUM DKO cells aggregate and form clusters which likely arise from an inability to escape cell-cell contacts

PUM DKO cells were observed to aggregate and form clusters after 2-3 days of culture (Figure 3.2B), in contrast to the even monolayer typical of WT T-REx-293 cells (Figure 3.2A). The slowed growth rate of PUM DKO cells may exacerbate this effect, since these cells took longer to generate similar cell volumes as WT. However, PUM DKO cells were unable to completely fill in the surface of a plate, even when cultured undisturbed for extended periods of time. This suggests that the phenotype observed cannot be due to differences in growth rates alone.

We hypothesized that this phenotype could arise through a few (not necessarily mutually exclusive) mechanisms: defects in motility, cytoskeletal structure/function, or adhesion, and that time-lapse imaging would be instrumental to observe the process and discern the mechanisms.

Time-lapse images of WT and PUMKO cells were taken and analyzed to compute several attributes that characterize and quantify the clumped appearance of these cells. Fractional area of cells within the image frame was measured, which quantifies the combined effect of cell morphology and cell growth rate. To uncouple the observed clumping phenotype with the decreased growth rate of PUM DKO cells, we also calculated the number of cell objects (defined as individual cells or clusters of contiguous cells) and the number of space objects (defined as a contiguous empty space surrounded by cells) per image throughout the time-lapse.

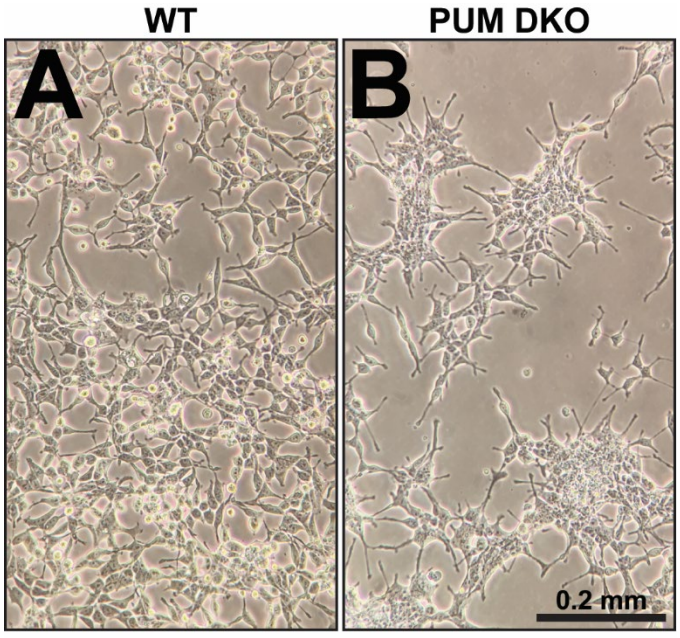



**Figure 3.2: PUM DKO cells are morphologically distinct from WT.** Representative images of WT (A) and PUM DKO (B) T-REx-293 cells. Images were taken 48 hours after plating. Bar represents 0.2 mm.

WT cells were able to form an even monolayer and occupy nearly all available area. Shortly after seeding, WT cells began growth in small and randomly dispersed clusters. In this phase of growth, the area occupied by cells was small (Figure 3.3A, image 1), but

composed of many cell objects (Figure 3.3B). At this time the space object number was near one, as most of the empty space is contiguous (Figure 3.3C). As WT cells divided and spread out, they quickly became evenly distributed and at ~30 hours the cells occupied about half of the total frame area (Figure 3.3A, image 2). At this phase, the even distribution of cells created one or a few contiguous cell objects (Figure 3.3B), and very characteristically contained many small space objects between cells (Figure 3.3C). Further growth and spreading lead to an increase in total cell area that approached the full area of the frame (Figure 3.3A, image 3) with one large cell object (Figure 3.3B) formed as the count of space objects rapidly decreased to a small number (Figure 3.3C).

In contrast, PUM DKO cells occupied a lower percent area and showed cell and space numbers distinct from WT. PUM DKO cells began growth similar in appearance to WT, with a low cell area (Figure 3.3A, image 4), many cell objects (Figure 3.3B), and one or few space objects (Figure 3.3C). As time passed, PUM DKO cells formed larger cell objects through their interactions with neighboring cells. Cells that contacted one another often remained stably attached, which lead to the formation of a higher number of discrete cell objects relative to WT at time interval 30-90 relative hours (Figure 3.3B, image 5). Examination of the time-lapse videos of individual frames clearly indicated that DKO cells that contact others during migration predominantly remained in attached clumps (Supplementary Video 2), while WT cells formed more transient contacts and could detach and continue individual movement (Supplementary Video 1). Over time, predominantly immobile cell clumps grew larger through cell divisions and absorption of/connection with other clumps. This resulted in the formation of an average of 10 distinct cell objects that

did not fully coalesce as in the WT (Figure 3.3B) and occupied an area ~35% lower than WT (Figure 3.3A, image 6). A gradual increase in space objects was formed by the gaps between cell clumps, which did not decrease as in the WT (Figure 3.3C).

Fractal dimension mean (FDM) and lacunarity are two image analysis metrics commonly used to characterize cells or tissues[36-39], and these criteria were used to further quantify differences between WT and PUM DKO cells. Fractal dimension is a measure of self-similarity within an image, where an increase in patterns or repetitive elements leads to a higher fractal dimension. The FDM of WT cells changed in accordance to the three states described above :1) the starting cell distribution of small and randomly dispersed clusters had a high FDM value, 2) evenly spaced cells that occupied half the area with many space objects had a low FDM, 3) confluent cells with little to no space objects had an intermediate FDM (Figure 3.3D). The FDM of PUM DKO cells started out the same as WT (as expected from observations) but then continually decreased over the course of 72 hours. PUM DKO cells slowly formed clumps and did not fully fill the frame area, a pattern that was similar to the early FDM decline in WT cells (Figure 3.3D). Lacunarity, a counterpart to fractal dimension, is a measure of the texture of an image, where images that have more hole/ gaps have a higher lacunarity. Lacunarity for WT cells reached its highest when cells were half confluent and contained many small space objects (Figure 3.3E). Lacunarity for PUM DKO cells increased over time as space objects formed and became larger (Figure 3.3E). In both FDM and lacunarity metrics, PDKO cells proceeded in a markedly distinct gradual trajectory, without a sharp minimum/maximum value observed in WT. Taken together, these results provide evidence that PUM DKO cells grow in a clumped distribution that is

distinct from WT cells, and that this effect can be decoupled from PUM-dependent changes in growth by observing time-lapse images that span the entire growth trajectory of the cell types.

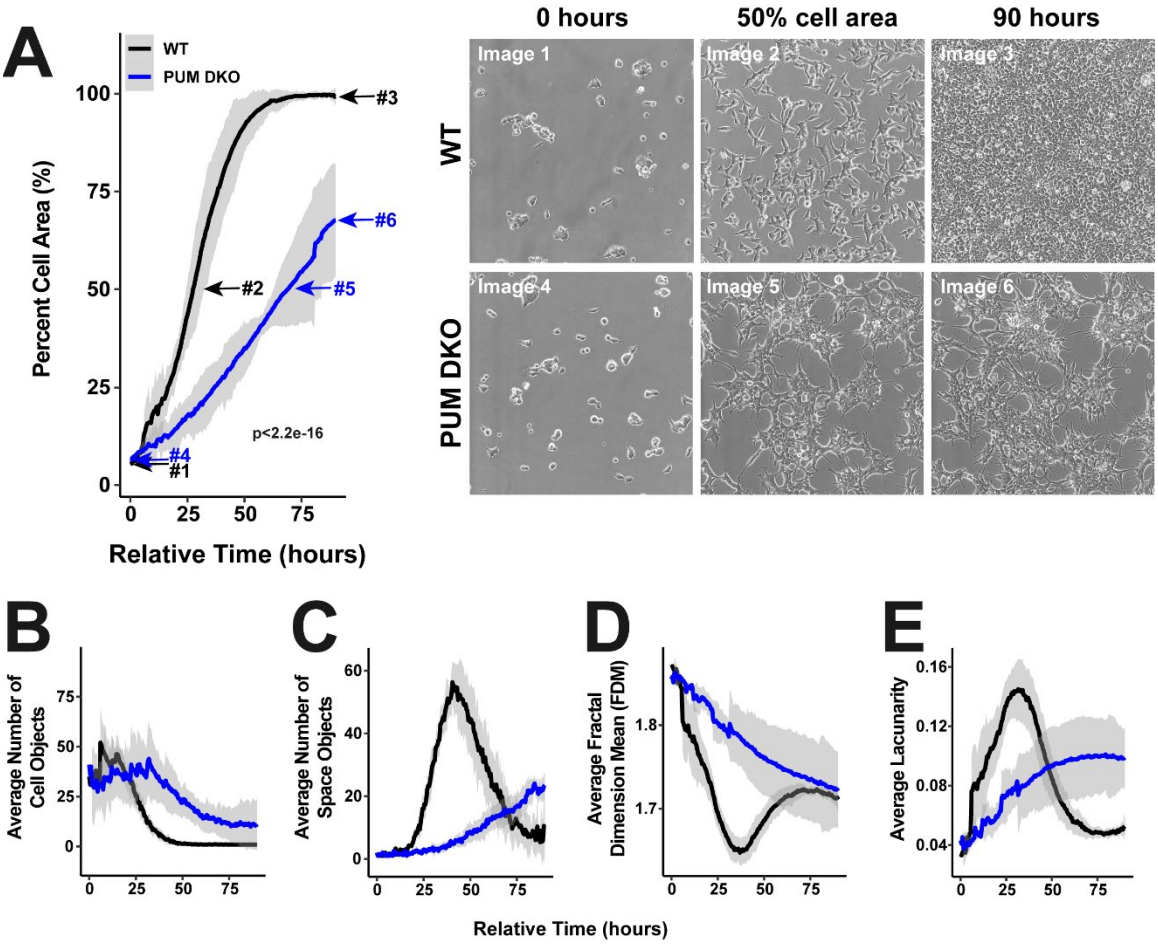



**Figure 3.3: PUM DKO cells aggregate and form clusters not typical of WT cells.** WT (black) and PUM DKO (blue) T-REx-293 average cell area (A), number of cell objects (B), number of space objects (C), fractal dimension mean (D), and lacunarity (E) over relative time. Representative images are shown in A. Measurements were calculated from six biological replicates. Shaded band represents two standard deviations from the mean. Statistical significance was determined by ANOVA F-test between a full multi-level model and a base model that ignored cell type, as described in Materials and Methods.

**Normal cell distribution is partially restored in both PUM1 and PUM2 rescues**

Reintroduction of either PUM1 or PUM2 abolished the clumped appearance we observed in PUM DKO cells and was accompanied by an increase in the average cell area over relative time (Figure 3.4A). Although this rescue was a highly significant increase in area compared to PUM DKO cells, rescue cells did not reach the full confluency of WT cells. Rescue cells displayed the same overall distribution pattern as WT cells (described above), but each phase took a longer time. Support of this was seen in the behavior of these cells: both PUM1 and PUM2 rescue cells started out similar to WT and PUM DKO cells (Figure 3.4, images 4 and 7), with a randomly dispersed low confluency pattern correlated with many cell objects (Figure 3.4B) and few space objects (Figure 3.4C). Over time rescue cells became evenly distributed and thus created an increasing number of space objects as time passed (Figure 3.4C). The number of space objects present in each rescue was higher than that of the GFP control and was reflective of the pattern we see in WT cells, although a sharp decrease in space objects towards the end of the time-lapse was not observed (Figure 3.3C). This is because both PUM1 and PUM2 rescues did not reach full confluency within the time frame of the experiment. Both PUM1 and PUM2 rescues are capable of forming an even monolayer (data not shown), and it is likely that a decrease in space objects would have been observed if measurements were carried out for a longer period of time (Figure 3.4D). Additionally, examination of time-lapse images of both PUM1 and PUM2 rescue cells showed that, unlike the DKO line, cells were able to detach after cell to cell contacts are made (Supplementary Videos 3 and 4). FDM and lacunarity measurements of both rescues displayed patterns similar to WT (Figure 3.4D and 3.4E). GFP expressing

PUM DKO cells behaved nearly identical to untransfected PUM DKO cells, which ruled out secondary effects of transfection on the measured attributes (Figure 3.4 (blue lines), images 1, 2, and 3, Supplementary Video 5). This provides evidence that PUM1 and PUM2 are directly responsible for the observed clumping phenotype.

Changes in cadherin expression are often correlated with changes in adhesion between neighboring cells. To test whether an increase in E-cadherin or N-cadherin could be responsible for the increased adhesion observed, levels of both proteins were measured by western blot (Supplementary Figure 3.4). E-cadherin was undetectable in both WT and PUM DKO cells, and there was no significant change in N-cadherin expression between cell types.

**Addition of a supplemented extracellular matrix is sufficient to rescue the clumped phenotype of PUM DKO cells**

To assess whether the clumping behavior is related to extracellular matrix components, a thin layer of extracellular matrix (Matrigel) was deposited on the attachment substrate. After cell plating, the clumped appearance of PUM DKO cells was lost (Figure 3.5). This observation could be due to many factors. One possibility is that PUM DKO cells are unable to properly produce or secrete an extracellular matrix, which leads to a change in their attachment and migration properties (although this is less supported by the observed motility of PUM DKO cells). Another possibility is that adhesion to extracellular matrix, or receptor binding to the factors within it, activates a signaling pathway that can compensate for loss of function in the DKO. Further experiments are needed to determine the mechanism.

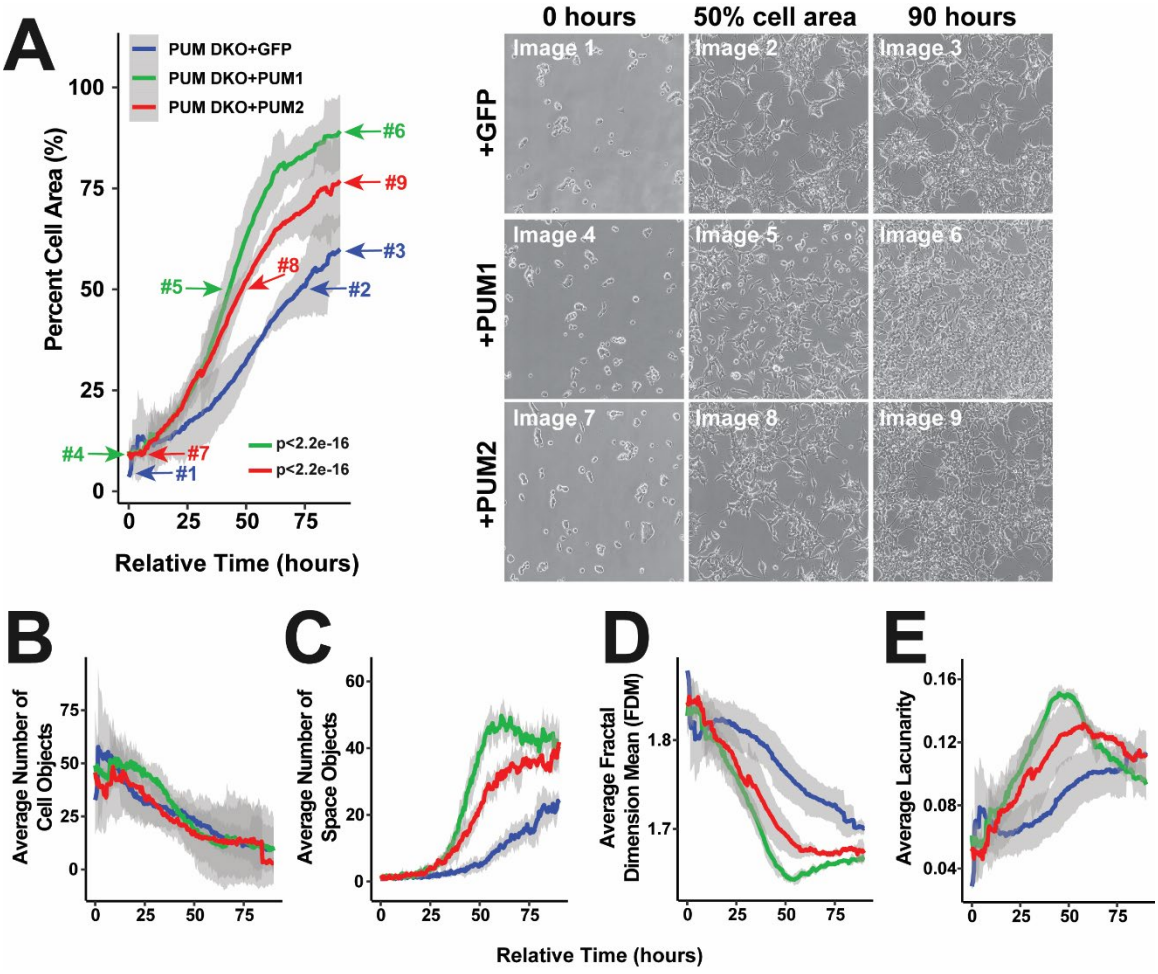



**Figure 3.4: DKO phenotype can be rescued by both PUM1 and PUM2.** Stable integrant populations of PUM1 (green), PUM2 (red), or GFP (blue) in PUM DKO T-REx-293 cells were analyzed for average cell area (A), number of cell objects (B), number of space objects (C), fractal dimension mean (D), and lacunarity (E) over relative time. Representative images are shown in A. Measurements were calculated from six biological replicates. Shaded band represents two standard deviations from the mean. Statistical significance was determined by ANOVA F-test between a full multi-level model and a base model that ignored cell type, as described in Materials and Methods.

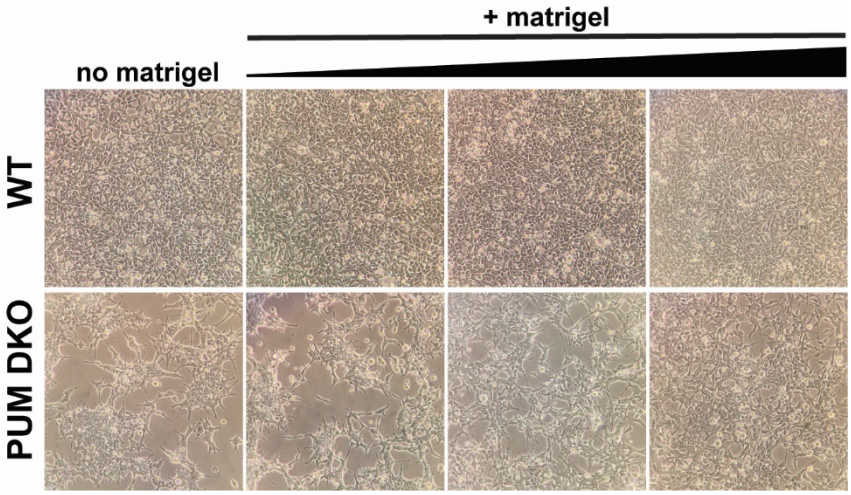



**Figure 3.5: The clumped distribution of PUM DKO cells is rescued by the addition of extracellular matrix.** WT and PUM DKO T-REx-293 cells were grown for 60 hours with varying dilutions of matrigel, ranging from 1/1000 to 1/20.

### PUM proteins affects the migration rate of cells in a clonal ring assay

To determine whether PUM proteins influence cell mobility, we performed a clonal ring migration assay. In this assay, cells were added to a cloning ring and given time to adhere to the plate at full confluency. The cloning ring was then removed, and collective cell movement into the surrounding free space was measured (Supplementary Videos 6 and 7). PUM DKO cells migrated at a significantly slower rate than that of WT cells (Figure 3.6A), and migration of cells was partially restored by the addition of either PUM1 or PUM2 (Figure 3.6B, Supplementary Videos 8 and 9). Interestingly, HCT116 cells, which also showed growth rate defects upon loss of PUM, did not show any difference in migration rate between WT and PUM DKO cells (Supplementary Figure 3.1). This provides evidence that the migration defect observed is separable from the growth rate defect.

Taken together with the time-lapse images, we hypothesized that the migration rate is partially affected by an increase in adhesion between PUM DKO cells, which hinders proper cell movement. Since clumping could be eliminated with the addition of Matrigel under normal culture conditions, we tested migration rates of WT and PUM DKO cells with and without the addition of Matrigel. The addition of Matrigel had no effect on the migration rates of WT and PUM DKO cells (Figure 3.6C) and did not rescue the observed clonal ring migration rates, which suggests that the DKO decrease in migration is not fully explained by the same mechanisms that lead to clumped distribution.

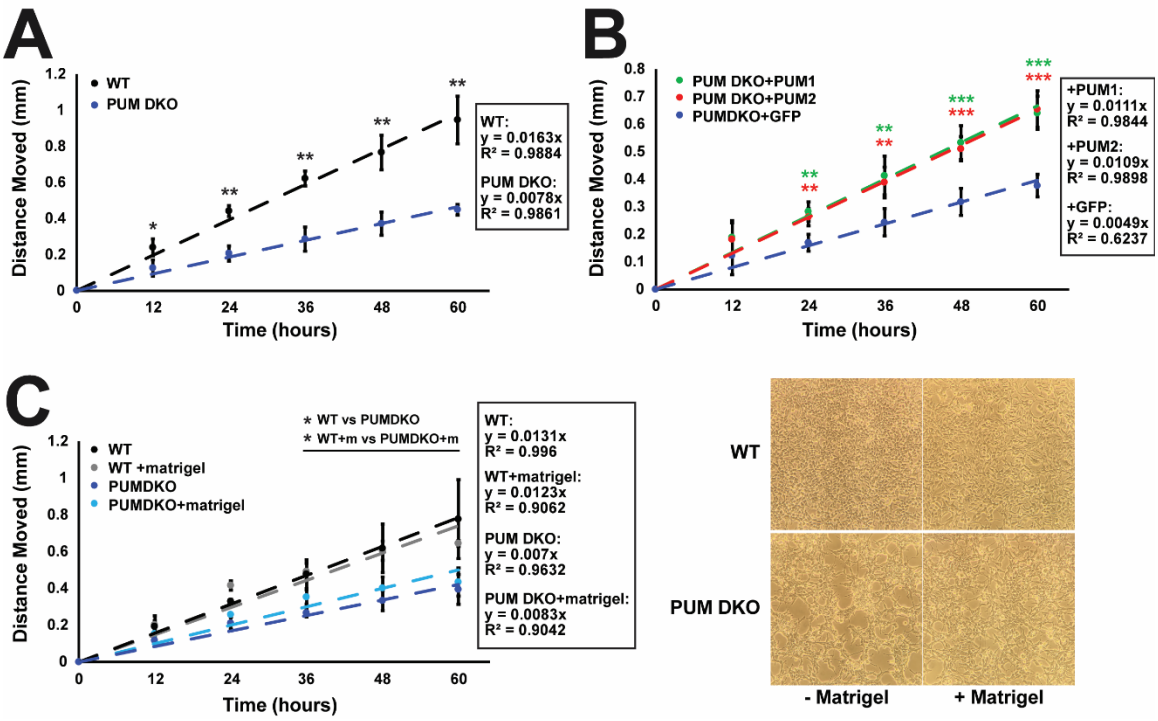



**Figure 3.6: PUM affects the migration rates of cells in a clonal ring assay.** (A) Migration rates of WT (black) and PUM DKO (blue) cells. (B) Migration mates of PUM DKO cells with the addition of PUM1 (green), PUM2 (red), and GFP (blue). (C) Migration rates of WT and PUM DKO cells with and without Matrigel. Matrigel was plated at a 1/50 dilution. Measurements were calculated over 3-5 biological replicates. Statistical significance was determined by Student's T-test. *= $p<0.05$, **= $p<0.01$, ***= $p<0.001$.

**PUM proteins affect actin cytoskeleton morphology**

Since a difference in migration was observed, we wanted to determine if any structural changes were present within the cytoskeleton. To this end, WT and PUM DKO cells were transfected with a Utrophin-RFP marker in order to label filamentous actin structures. Fluorescence and brightfield images of mainly individually situated (non-touching) live cells were scored based on their major actin cytoskeletal structure related to locomotion: lamellipodia (flat, usually broad, plate-like extensions), filopodia (long, slender extensions), or other (does not fit either other description). Representative images are shown in Figure 3.7A. To eliminate predisposed bias, cell genotypes were blinded to the scorer. WT cells had a strong preference to be filopodia dominant, while PUM DKO cells were observed to have an even amount of filopodia and lamellipodia dominant cells, a difference with statistical significance (Figure 3.7B). This result provides evidence that PUM proteins impact the cytoskeletal behavior of T-REx-293 cells, which could underlie the observed defects in motility.

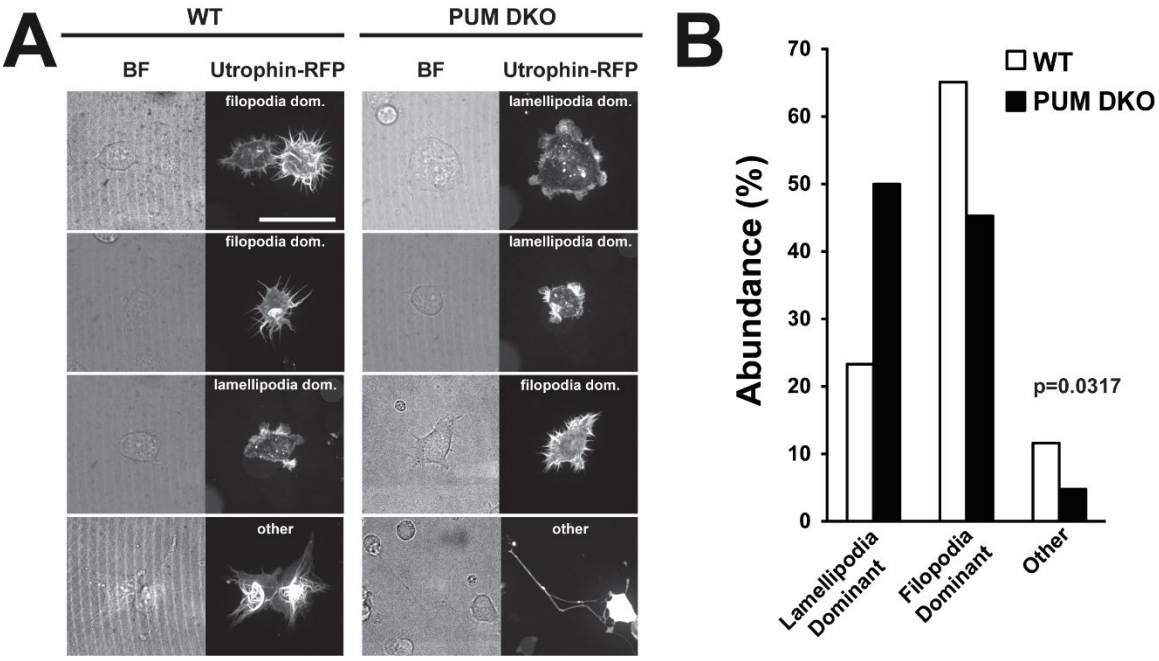



**Figure 3.7: PUM DKO cells show changes in actin morphology.** (A) Representative brightfield (BF) and RFP images of WT and PUM DKO T-REx-293 cells. Scored category is noted in each image. White bar represents 50 µm. (B) Percent abundance of cells categorized as lamellipodia dominant, filopodia dominant, or other. Cell images were scored blindly. Distributions were compared using a Chi-squared test.

## RNA-sequencing reveals a large set of genes affected by the PUM proteins, and an enrichment in taxis and extracellular protein processes

RNA-sequencing was performed in order to determine genes and pathways that may be disrupted in the PUM DKO cells. To do this, three biological replicates for WT, PUM DKO, and the three rescue constructs were generated. There were a total of 1143 differentially expressed genes between WT and PUM DKO (at a 2-fold and <5% FDR cutoff), 751 of which genes were upregulated and 392 were downregulated (Figure 3.8A). The fact that more genes were upregulated upon PUM DKO agrees with PUMs role as a repressive protein. However, approximately one third of genes were downregulated in PUM DKO cells. This fraction is likely in part is due to secondary effects, but some of the effect may be attributed to a stabilizing role of the PUM proteins. The addition of PUM1

lead to 696 and 127 differentially expressed genes when compared to PUM DKO cells and the GFP control, respectively. PUM2 addition showed a much smaller effect on gene expression than PUM1, with 28 and 9 differentially expressed genes compared to PUM DKO and the GFP control, respectively. RNA-sequencing for three biological replicates of HCT116 WT and PUM DKO cells were also generated and are summarized in Supplementary Figure 3.1.

Intriguingly, gene ontology analysis for both cellular component and biological process showed enrichment in categories related to adhesion and migration. Extracellular matrix and cell surface localized proteins were the top scoring cellular component categories within genes that were upregulated in PDKO vs WT cells (Figure 3.8B), and the same gene set showed biological process enrichment in genes involving taxis and extracellular organization as top scoring categories (Figure 3.8C). In many cases, the direction of mRNA level change for an enriched category was coherent across comparisons, i.e. when a category was enriched among genes upregulated in PUM DKO cells compared to WT, the same category would be enriched among genes that are downregulated in rescue constructs compared to PUM DKO cells. This indicates that reintroduction of PUM into the KO setting restores the levels of many genes toward their WT values. To identify candidate genes that may be tied to the motility and adhesion phenotypes we observed, the RNA-sequencing dataset was filtered for genes that either went up in the DKO compared to WT and then went down in both rescues compared to PDKO or went down in the DKO compared to WT and then went up in both rescues compared to PDKO, and did not show significant change in PUM DKO cells compared to the GFP control. This selection lead to

a short list of 21 candidate genes that may be tied to the phenotypes we observe (Table 3.2). Of these genes, eight contained PUM motifs (defined as 5'-UGUANAUW -3') and were also bound by either PUM1 or PUM2 in a CLIP-sequencing dataset generated in our lab[10]. Of those eight, six genes were also bound in independently generated PAR-CLIP and/or RIP-Chip datasets[12,40,41] . These genes are strong candidates for direct PUM regulation. Further testing is needed to confirm whether these candidate genes are true PUM targets and responsible for the adhesion and migration defects observed.

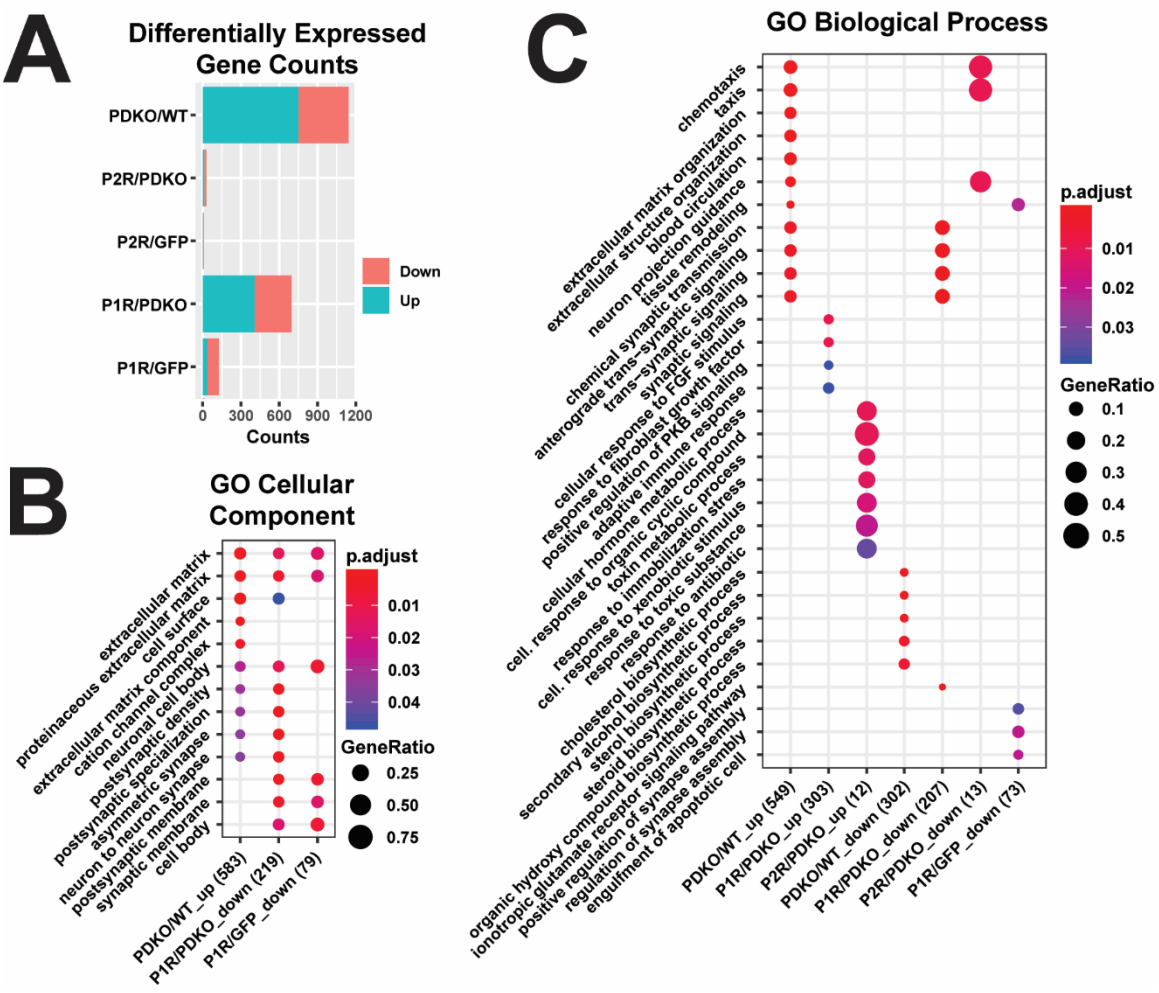



**Figure 3.8: RNA-sequencing reveals a large set of genes affects by the PUM proteins, and an enrichment in taxis and extracellular protein processes. (A)** Number of differentially expressed genes for WT, PUM DKO, and PUM rescue constructs. Gene ontology for cellular component (B) and biological process (C).

**Table 3.2: List of candidate genes that may be involved in PUM-dependent cell adhesion phenotype.**

| | PUM DKO vs WT | | PUM DKO+PUM1 vs PUM DKO | | PUM DKO+PUM2 vs PUM DKO | | PUM DKO+GFP vs PUM DKO | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SYMBOL** | **$Log_2FC$** | **FDR** | **$Log_2FC$** | **FDR** | **$Log_2FC$** | **FDR** | **$Log_2FC$** | **FDR** | **Bound in T-REx-293 CLIP** | **PUM motif in 3'UTR** | **Bound in other RIP-Chip/ PAR-CLIP datsets** | **DE in other RNA-seq** |
| CCDC12 | 1.98 | **1.3E-22** | -1.16 | **2.5E-07** | -0.76 | **2.3E-02** | 0.02 | 9.93E-01 | YES | YES | YES | YES |
| FZD8 | 0.93 | **2.4E-08** | -0.82 | **4.8E-06** | -0.56 | **4.1E-02** | 0.05 | 9.82E-01 | YES | YES | YES | YES |
| ISCU | 1.02 | **1.6E-11** | -0.65 | **1.7E-04** | -0.61 | **5.6E-03** | -0.28 | 8.03E-01 | YES | YES | YES | YES |
| NECTIN4 | 3.12 | **1.5E-09** | -2.01 | **8.2E-05** | -1.57 | **2.3E-02** | -0.40 | 9.33E-01 | YES | YES | YES | - |
| PHAX | -0.87 | **3.4E-05** | 0.74 | **1.5E-03** | 0.76 | **1.0E-02** | 0.40 | 7.77E-01 | YES | YES | YES | NO |
| PTPRA | 1.42 | **9.8E-26** | -0.90 | **1.1E-09** | -0.48 | **5.0E-02** | -0.27 | 8.03E-01 | YES | YES | YES | NO |
| CRISPLD1 | 1.19 | **2.9E-02** | -1.19 | **3.8E-02** | -1.59 | **3.7E-02** | -1.00 | 6.77E-01 | YES | YES | NO | YES |
| SMIM29 | 1.21 | **4.1E-05** | -1.48 | **8.5E-07** | -1.03 | **1.5E-02** | -0.26 | 9.33E-01 | YES | YES | - | - |
| LNX2 | -0.64 | **3.3E-02** | 0.92 | **1.5E-03** | 0.84 | **4.5E-02** | 0.53 | 7.11E-01 | YES | NO | YES | NO |
| PLAT | 1.56 | **2.0E-08** | -1.59 | **4.4E-08** | -0.97 | **2.9E-02** | -0.09 | 9.78E-01 | YES | NO | YES | YES |
| SDHA | 0.61 | **7.1E-04** | -0.63 | **8.6E-04** | -0.56 | **4.1E-02** | -0.29 | 8.20E-01 | YES | NO | YES | NO |
| TFPI | -1.04 | **1.9E-02** | 1.75 | **1.9E-05** | 1.48 | **6.3E-03** | 1.17 | 2.43E-01 | NO | NO | YES | NO |
| APOBEC3B | -1.10 | **1.0E-02** | 1.87 | **1.9E-06** | 1.46 | **5.6E-03** | 1.30 | 5.30E-02 | NO | NO | NO | NO |
| DPYSL4 | 1.08 | **4.1E-04** | -1.95 | **2.6E-10** | -1.12 | **5.6E-03** | -0.67 | 5.55E-01 | NO | NO | NO | NO |
| RAP1GAP | 1.04 | **2.2E-04** | -1.25 | **1.7E-05** | -1.01 | **8.2E-03** | -0.45 | 8.14E-01 | NO | NO | NO | NO |
| SCAMP5 | 1.44 | **1.0E-05** | -1.87 | **1.3E-08** | -1.18 | **1.1E-02** | -0.92 | 3.08E-01 | NO | NO | NO | YES |
| TAF7 | -0.39 | **4.2E-02** | 0.66 | **2.5E-04** | 0.59 | **1.5E-02** | 0.40 | 5.40E-01 | NO | NO | NO | NO |
| ADGRB1 | 1.31 | **1.4E-08** | -1.31 | **6.0E-08** | -0.79 | **3.5E-02** | -0.30 | 9.14E-01 | NO | NO | - | - |
| CCDC181 | -0.97 | **1.2E-02** | 1.54 | **2.4E-05** | 1.19 | **2.2E-02** | 0.77 | 6.27E-01 | NO | NO | - | - |
| DRAXIN | 2.07 | **1.2E-08** | -1.68 | **7.3E-06** | -1.47 | **9.0E-04** | -0.69 | 6.64E-01 | NO | NO | - | - |

## Discussion

We have phenotypically characterized the effect of PUM double knock out in T-REx-293 cells and have shown that loss of PUM leads to defects in adhesion and motility. PUM DKO cells predominantly remained attached once contact is made, whereas WT cells formed transient interactions and were able to detach and continue individual movement. This leads us to believe that the clumping phenotype observed is due to defects in cell adhesion. Although we observed no differences in expression levels of E-cadherin and N-cadherin, this does not rule out that changes in their localization to the plasma membrane and/or protein modifications could occur in PUM DKO cells, affecting overall cell adhesion. We have also not ruled out proteins that contribute to or regulate cell junctions. Similarly, we have not tested all cadherins present in T-REx-293 cells, although the remaining cadherins are unconventional and are expressed at low levels in this cell type.

Co-culture experiments showed that addition of RFP-labeled WT cells to GFP-labeled PUM DKO cells was sufficient to rescue the clumping phenotype (Supplementary Figure 3.6), and PUM DKO cells in the presence of WT cells seemed more able to break cell-cell adhesions. In addition, PUM DKO cells did not segregate into distinct spatial areas from WT cells, which lends evidence against differences in cadherin expression. This, along with the Matrigel experiments, lead us to believe that misregulation of a secreted protein may contribute to the phenotypes observed. It is important to note that addition of extracellular matrix can only rescue the clumping phenotype and does not rescue the migration rate defect. This suggests that the two phenotypes are separable and may be rooted in misregulation of different factors.

Changes in actin morphology point to changes in activity of the Rho family of GTPases. These pathways, along with regulating actin structure, are known to regulate cell motility and cell adhesion. Because of this, GTPase activating proteins and guanine exchange factors, along with the GTPases themselves, are promising targets. PUM proteins regulation of Rho GTPases is supported by other studies which show an enrichment in transcripts encoding GTPases as being targets of PUM.

PUM1 displays an increased ability to rescue growth and adhesion phenotypes. This is supported by the results of the RNA-sequencing data, where more genes were differentially expressed upon addition of PUM1 that PUM2. Western blots confirm that both PUM1 and PUM2 were rescued to approximately the same level, so this effect is unlikely to be due to protein levels (Figure 3.1B). It has been noted previously that PUM1 and PUM2 have some non-overlapping functions. PUM1 contains a longer N-terminal domain which may direct distinct behaviors, possibly through interactions with other proteins. Single rescues of each PUM allow for the identification of protein specific behaviors of each, and the observation that neither PUM protein is sufficient for compete rescue, lends evidence to non-overlapping functions of the PUM proteins. However, this effect may also be explained by dosage effects[32], since the total PUM expression level is less in each rescue than in WT cells.

Differences in growth rate between WT and PUM DKO cells has already been observed, and further validates known effects of PUM proteins. This phenotype is tied to PUMs regulation of cell cycle machinery and regulatory proteins, as well as its role in genome stability. In our study, we used cell counting by hemocytometer to measure growth over

time. With more precise measurements, more specific aspects of the rate equation may become significantly different from each other and could provide insight into mechanisms of growth that differ between conditions. PUM2 rescue growth rates may also be statistically significant from the GFP rescue. Again, the fact that each rescue does not restore growth to WT levels suggests either non-redundant roles of the PUM proteins or a protein dosage effect.

**Candidates from an integrative analysis of RNA sequencing, PUM CLIP data and motif detection**

When selecting candidates for transmitting PUM effects from the RNA-sequencing data, we required that transcript abundance be significantly changing in the WT/PUM DKO comparison and be restored in both PUM1 and PUM2 rescues. However, experiments point to PUM1 playing a larger role in adhesion and migration, and so loosening the requirement of PUM2 to restore the transcript may increase the list of valid candidates. We also compared top scoring candidates to a pre-existing CLIP-sequencing dataset from our lab in order to identify direct targets of PUM. This is overall a good indicator of true PUM binding, although binding does exist at partial PRE sites and even at sites with no obvious PRE. A modeling in human cells demonstrates that additional sites can be recognized through nucleotide flipping[42]. There are also instances of PREs not showing bound PUM proteins, and it is demonstrated that RNA modifications can lead to weakened binding of the PUM proteins[43]. A few candidates decrease in expression upon PUM loss, suggesting a stabilizing role of PUM. This can be due to secondary effects, although examples of a stabilizing role of PUM exist across organisms[10,29,44]. A meta-analysis of both CLIP-

sequencing and RNA-sequencing datasets in HEK cells has also revealed a population of transcripts that are both bound by PUM and decrease in expression upon PUM knock down[11]. Because of this, we included both repressed and stabilized transcripts in our final candidate list.

Within the RNA-sequencing data, comparisons between rescues and the GFP control (rather than PUM DKO cells) returns fewer differentially expressed genes. This is likely in part because the PUM1 and PUM2 rescue, and the control GFP-expressing cell lines were derived in parallel, and all underwent an additional round of selection with the integration of the relevant plasmid. Principal component analysis also showed that there was an increased amount of variability between replicates of PUM1, PUM2, and GFP expressing cells (Supplementary Figure 3.5). The variance here seems to come from the replicate two sample preparations, which were all prepared in parallel. This variability introduced some noise into the data, and some differentially expressed genes may have fallen out of significance because of this.

Of our top scoring candidates, a couple stand out as promising candidates. Frizzled-8 (FZD8) is a non-canonical Wnt protein receptor[45], and non-canonical Wnt signaling is known to play a role in the regulation of cell adhesion and migration[46]. NECTIN-4 belongs to a class of proteins involved in cell-cell adhesion, and expression of NECTIN-4 on the surface of ovarian cancer cells increases adhesion, although its expression is also associated with increased migration in a scratch assay[47].

Adhesion and migration play important roles in coordinating many cell behaviors and processes. Proper control of cell-cell adhesion and migration are essential for cell

organization during development. In addition, cancer cells are known to increase malignancy through decreasing contacts with neighboring cells and in some cases upregulating motility genes[48,49]. Understanding the role of PUM proteins in in these processes is essential for providing better models for its function in developmental processes, as well as understanding how changes in PUM expression can influence cancer.

## Materials and Methods

### Cell culture

T-REx-293 cells were obtained from Invitrogen. HCT116 cells were a gift[32]. PUM double knockout cells were generated as described previously[10]. T-REx-293 cells were grown in DMEM (Corning) with 10% fetal bovine serum (Corning) and 10 units/mL of penicillin/streptomycin (Gibco) at 37°C with 5% $CO_2$. HCT116 cells were grown in McCoy's 5A (Iwakata & Grace Modification) media (Corning) with 10% fetal bovine serum (Corning) and 10 units/mL of penicillin/streptomycin (Gibco) at 37°C with 5% $CO_2$. For the addition of Matrigel Matrix (Corning), 1 mL total volume of ice cold Matrigel was added to a 6-well and incubated for 1 hour at room temperature. Excess solution was removed just prior to plating. Appropriate ice-cold media was used to dilute Matrigel from 1/1000 to 1/20.

### Rescue cell lines

Overexpression plasmids for PUM1 (pLX302-PUM1), PUM2(pLX302-PUM2), and GFP (pLX302-GFP) were a gift[35]. TransIT-LT1 reagent (Mirus) was used per manufacturer's instruction to add 1 µg of plasmid to PUM double knockout T-REx-293 cells seeded in 6-well plates at ~70% confluency. 48 hours after transfection, cells were selected with 1 µg/mL puromycin for at least 7 days. PUM1 and PUM2 expression was confirmed by western blot using a goat anti-PUM1 antibody (Bethyl, A300-201A) and a rabbit anti-PUM2 (Bethyl, A300-202A).

**Cell growth measurements**

Cells from an 80-90% confluent plate were trypsinized, counted on a hemocytometer and plated at an initial density of 325,000 cells in multiple wells of 6-well plates. At regular intervals, individual wells were harvested for counting on a hemocytometer, averaging 2-3 1 mm x 1 mm hemocytometer squares for each biological replicate. A total of 3 biological replicates were measured. Growth parameters were derived from standard logistic growth nls model fits in R. Statistical significance of differences in model fits between cell types were determined by pairwise F-test ANOVA comparisons of nested models that incorporate or ignore the cell type, as described[50].

**Image analysis of time-lapse microscopy images**

Cells from an 80-90% confluent plate were trypsinized, counted on a hemocytometer and plated at an initial density of 325,000 cells in a 6-well plate 24 hours prior to imaging. Phase contrast images were collected with a Biostation CT over 72 hours every 10-15 minutes. A total of 6 biological replicates for each cell type, consisting of 2-3 separately imaged regions ("frames", technical replicates) per each cell type / biological replicate well were collected, and images at coinciding 30-minute intervals were used for further analysis. Area occupied by cells and the number of cell and space objects were computed by the available functions in CL Quant (Nikon). Fractal dimension mean and lacunarity were computed by Matlab scripts as described[39]. All data was further analyzed and averaged in R. Individual biological and technical replicates exhibited trajectories of computed metrics (cell area, object number, FDM, lacunarity) over time that were very stereotypical within a given cell type but were shifted relative to each other along the x (time) axis

(Supplementary Figure 3.2A). The shifts resulted from stochastic variability in local seeded density between the imaged frames, and a variable cell recovery lag phase after the seeding of cells and prior to the onset of active growth and motility. To eliminate this variability, the time axes of individual frames within a given cell type were shifted such that the cell areas in the frames were aligned to each other with maximal overlap using the dtw R package with a rigid step pattern (Supplementary Figure 3.2B) and averaged within each cell type. This operation thus re-aligned the time axis of each technical replicate based on having the same growth stage (cell area). To compare between cell types, the time axes of averaged traces were aligned to each other over the initial 10% of their respective growth curves, to be able to analyze subsequent growth from a starting point of equivalent cell area (Figure 3.3A and Figure 3.4A). The newly established time axes are indicated as relative time. Computed object counts, FDM and lacunarity exhibited similar time shifts for individual frames, and the trajectories showed strong agreement when aligned by the relative time based on the cell area (Supplementary Figure 3.2C and 3.2D) and were analyzed using these alignments. To test for statistical significance of the differences in cell area, we built multi-level statistical models with a cubic orthogonal polynomial fit of each frame's cell area, including fixed effects for cell type with all possible interactions with the polynomial coefficients, and random effects for the linear and quadratic polynomial coefficients (models with random effects for all three polynomial coefficients did not converge in fitting). The analysis follows that as described[51]. Statistical significance was determined by ANOVA F-test between the full model and a base model that ignored cell type.

**Clonal ring migration assays**

Cells from an 80-90% confluent plate were trypsinized, counted on a hemocytometer and plated at an initial density of 800,000 or 1,000,000 for T-REx-293 and HCT116 cells, respectively, inside a 6 mm I.D. cloning ring within a 6-well. Cells were given 4-6 hours to attach to the plate before the cloning ring was removed. Images were taken every 12 hours for 60 hours to track collective migration of cells. Images were processed, aligned, and measured in ImageJ. For migration assays supplemented with Matrigel, a 1/50 dilution was used (protocol as described above).

**Fluorescence Microscopy**

TransIT-LT1 reagent (Mirus) was used per manufacturer's instruction to add 1 µg of the mRFP-UtrCH plasmid (Addgene #26739) to both WT and PUM DKO T-REx-293 cells seeded in 6-well plates at ~70% confluency. Cells were trypsinized and resuspended during transfection in order to increase transfection efficiency. 48 hours after transfection, cells were trypsinized and plated onto NaOH (2M for 2 hours) and poly-lysine (0.5 mg/mL on shaker for 1 hour) treated glass-bottom plates. Live cells were imaged 24 hours post seeding using a custom-built spinning disk confocal microscope (Solamere Technology) with a Yokagawa W1 spinning disk (Yokagawa), EM-CCD camera (Hamamatsu 9100c), and a Nikon Eclipse TE (Nikon) inverted stand. A 60× water immersion lens (1.2 NA) was used with perfluorcarbon immersion liquid (RIAAA-678, Cargille). The stage is fully motorized and controlled by Micromanager software (www.micromanager.org) with ASI Peizo (300-µm range) and a 3 axis DC servo motor controller. Solid-state lasers (Obis from

40 to 100 mW) and standard emission filters (Chroma Technology) were used. A 561 laser with emission filter 620/60 was used.

**RNA-seq library preparation and analysis**

For each cell type, three biological replicates were collected and processed separately. Cells were cultured to 50% confluency in a 10cm plate. Total RNA was extracted with Ribozol, and libraries were prepared using the NEB mRNA magnetic isolation module (E7490S), NEBNext Ultra RNA Library prep kit (E7420L), and NEB Multiplex Oligos for Illumina Index Primer Sets 1 (E7335S) and 2 (E75500S) and sequenced on an Illumina NextSeq instrument. Further processing was done in R using the systemPipeR[52] workflow: reads were aligned to the GRCh38 genome assembly using HISAT2 and annotated with Gencode v27 annotations. Differentially expressed genes were identified using DESeq2, and gene ontology enrichment was determined using the clusterProfiler package.

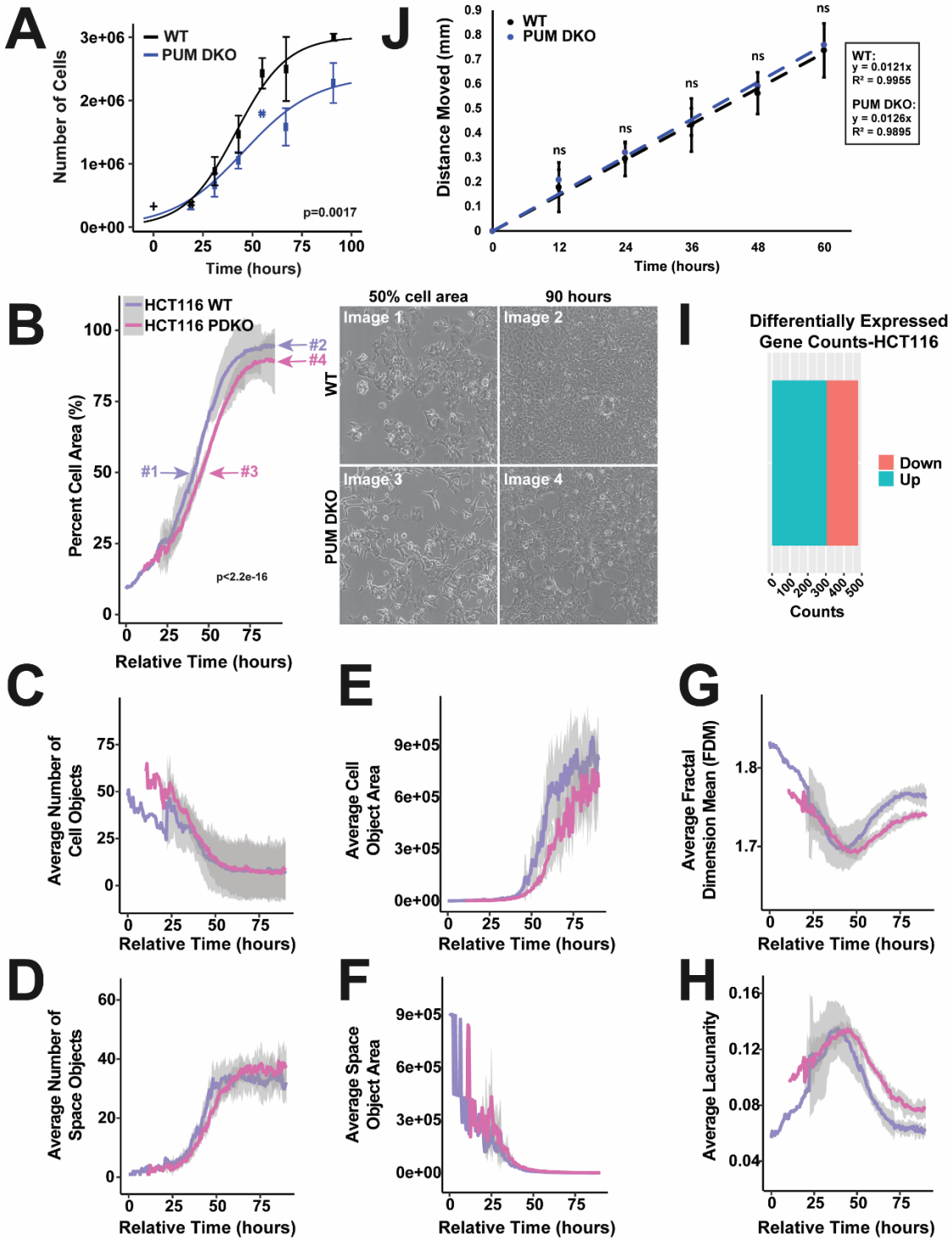
**Candidate western blots**

Antibodies used for western blots are as follows: mouse anti-E-cadherin (BD Transduction Laboratories, 610181 and Santa Cruz, sc-8426), mouse anti-N-cadherin (Santa Cruz, sc-393933), mouse anti-ephrin-B1 (Santa Cruz, sc-515264), mouse anti-Wnt-5a (Santa Cruz, sc-365370), mouse anti-c-Jun (Santa Cruz, sc-74543), mouse anti-GSK-3a (Santa Cruz, sc-5264), and mouse anti-BPIX (Santa Cruz, sc-393184). T-REx-293 and HCT116 cells were collected at ~70% confluency and run into a SDS-polyacrylamide gel. Gels were transferred to a nitrocellulose membrane, blocked, and incubated with primary antibody overnight at 4°C.

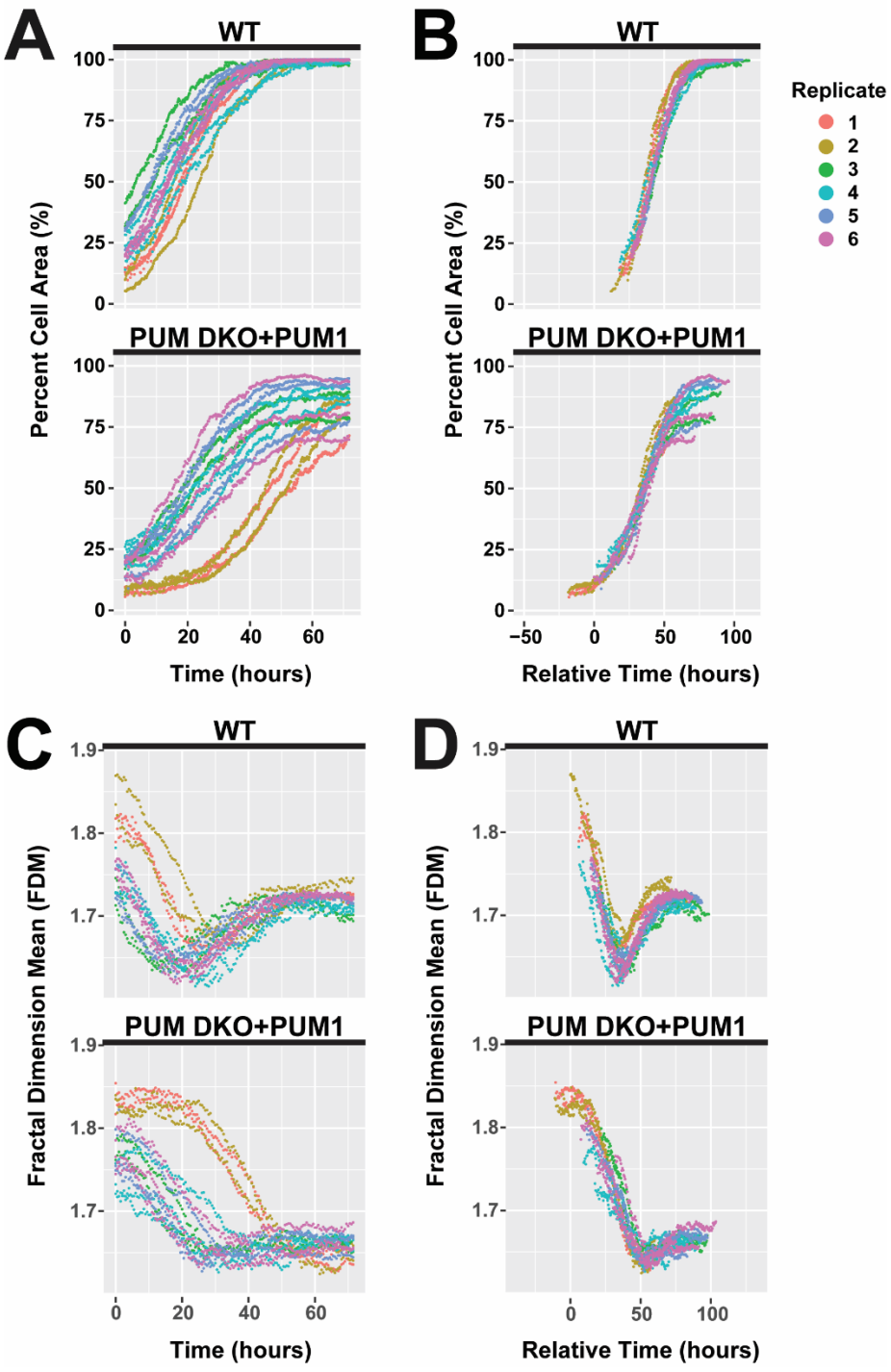
**Mixing Experiments**

The pCMV DsRed-Express2 (Clontech 632539) expression plasmid was stably transfected into wildtype T-REx-293 cells by same protocol described above and was selected using Neomycin at 500 μg/mL for seven days. The pMSCV-PIG (Addgene #21654) expression plasmid was stably transfected into wildtype 293T cells by the same protocol described above and was selected using Puromycin at 1 μg/mL for seven days. Generation of T-Rex-293 PUM DKO+GFP cells is described above. Cells were co-cultured at a ratio of 1 to 1 and a total cell volume of 325,000 cells per 6-well. Time-lapse images were collected as described above.

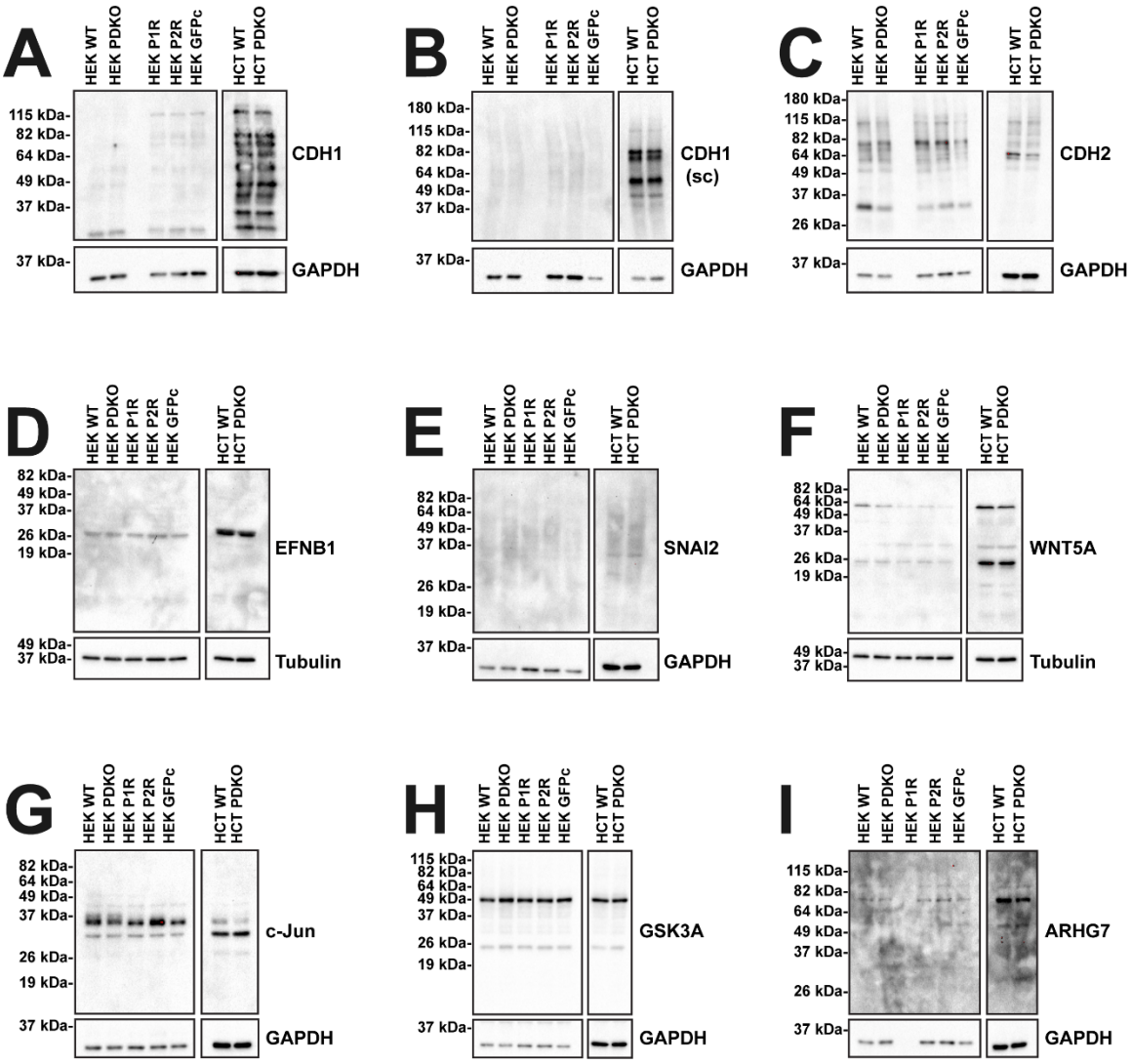
## Supplementary Figures



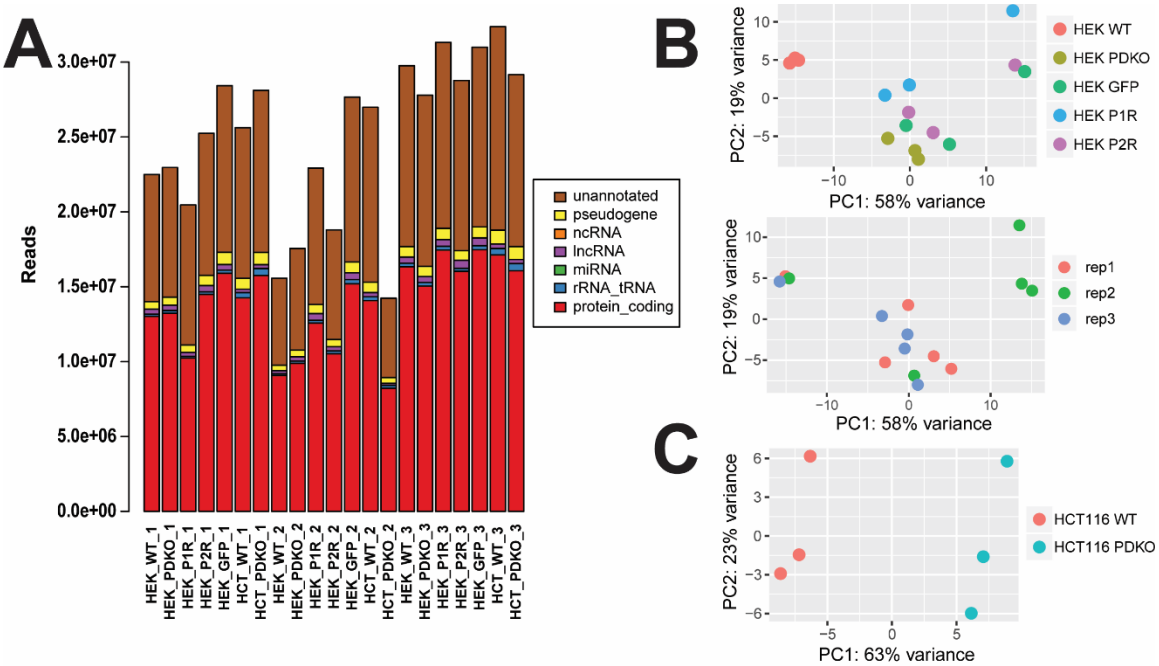
**Supplementary Figure 3.1: HCT116 PUM DKO cells show growth defects, but not strong adhesion and migration defects.** (A) Growth rates of WT (purple) and PUM DKO (pink) HCT116 cells. Growth measurements were calculated from three biological replicates, with error bars representing standard error of the mean. For the purposes of plotting, the time points are grouped for each cell type / replicate group (since the data was collected at slightly different timepoints for each replicate), generating error bars along the x axis direction (standard error of the mean). Logistic growth model fits of the data were compared using ANOVA F-test to determine statistical significance. (B-H) WT (purple) and PUM DKO (pink) average cell area (B), number of cell objects (C), number of space objects (D), cell object area (E), space object area (F), Fractal Dimension Mean (G), and Lacunarity (H) over relative time. Representative images are shown in B. Measurements were calculated from six biological replicates. Shaded band represents two standard deviations from the mean. Statistical significance was determined by ANOVA F-test between a full multi-level model and a base model that ignored cell type, as described in Materials and Methods. (I) Number of differentially expressed genes between WT and PUM DKO cells. (J) Migration rates of WT (black) and PUM DKO (blue) cells. Measurements were calculated over 3 biological replicates. Statistical significance was determined by Student's T-test.



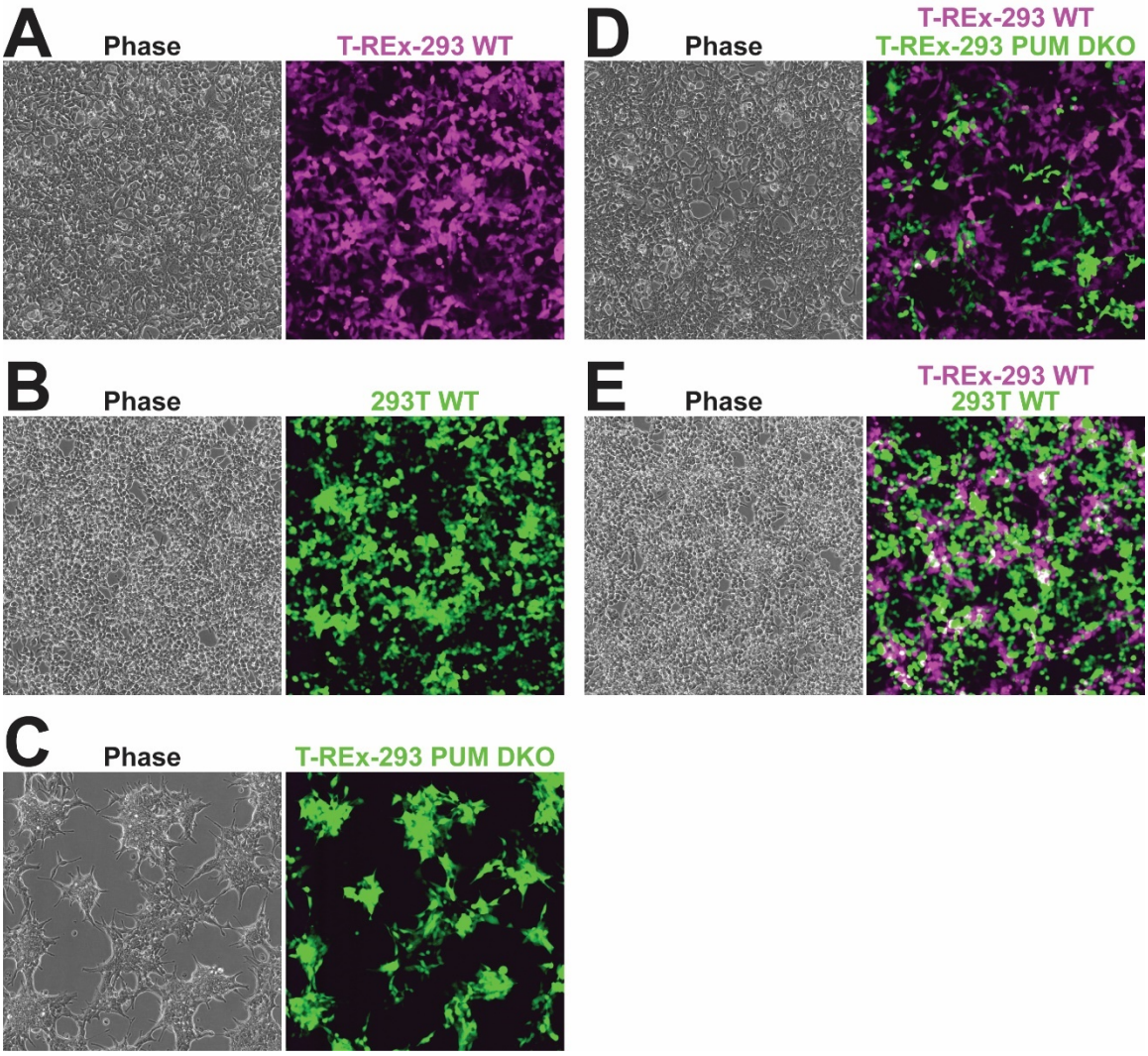

**Supplementary Figure 3.2: Example Time-lapse image analysis plots before and after curve alignments.** (A,B) Percent cell area plots for WT and PUM DKO+PUM1 T-REx-293 cells before (A) and after (B) alignment. (C,D) Fractal dimension mean plots for WT and PUM DKO+PUM1 T-REx-293 cells before (C) and after (D) alignment.




**Supplementary Figure 3.3: The PUM DKO adhesion phenotype is not caused by changes in cadherin expression levels, or levels of select candidate genes.** Western blots across WT, PUM DKO, and rescue T-REx-293 cells for E-cadherin (A, B), N-cadherin (C), Ephrin B1 (D), Snail2 (E), Wnt5A (F), Snai1 (G), c-Jun (H), GSK-3$\alpha$ (H), and Arhgef7 (I).




**Supplementary Figure 3.4: RNA-sequencing characteristics and quality.** (A) Stacked bar plots of each sample replicate, where reads are categorized by annotation category. (B) Principal component analysis for T-REx-293 cells color coded by cell type (top) and replicate number (bottom). (C) Principal component analysis for HCT116 cells color coded by cell type.




**Supplementary Figure 3.5: Co-culture of WT and PUM DKO cells rescues the increased cell adhesion phenotypes.** Phase contrast and fluorescence images of T-REx-293 WT only (A), 293T WT only (B), T-REx-293 PUM DKO only (C), co-cultured T-REx-293 WT and PUM DKO (D), and co-cultured T-REx-293 WT and 293T WT cells. Images were collected after 72hours of culture.

## Supplementary Video Legends:

**Supplementary Video 1: Time-lapse video of WT T-REx-293 cells.** Frames correspond to 15-minute intervals over a 72-hour period.

**Supplementary Video 2: Time-lapse video of PUM DKO T-REx-293 cells.** Frames correspond to 15-minute intervals over a 72-hour period.

**Supplementary Video 3: Time-lapse video of PUM DKO T-REx-293 cells stably expressing PUM1.** Frames correspond to 15-minute intervals over a 72-hour period.

**Supplementary Video 4: Time-lapse video of PUM DKO T-REx-293 cells stably expressing PUM2.** Frames correspond to 15-minute intervals over a 72-hour period.

**Supplementary Video 5: Time-lapse video of PUM DKO T-REx-293 cells stably expressing GFP.** Frames correspond to 15-minute intervals over a 72-hour period.

**Supplementary Video 6: Video of WT T-REx-293 cell migration in clonal ring assay.** Frames corresponds to 0, 12, 24, 36, 48, and 60 hour timepoints. Black lines are present for image alignment.

**Supplementary Video 7: Video of PUM DKO T-REx-293 cell migration in clonal ring assay.** Frames corresponds to 0, 12, 24, 36, 48, and 60 hour timepoints. Black lines are present for image alignment.

**Supplementary Video 8: Video of PUM DKO+PUM1 T-REx-293 cell migration in clonal ring assay.** Frames corresponds to 0, 12, 24, 36, 48, and 60 hour timepoints. Black lines are present for image alignment.

**Supplementary Video 9: Video of PUM DKO+PUM2 T-REx-293 cell migration in clonal ring assay.** Frames corresponds to 0, 12, 24, 36, 48, and 60 hour timepoints. Black lines are present for image alignment.

**Supplementary Video 10: Video of PUM DKO+GFP T-REx-293 cell migration in clonal ring assay.** Frames corresponds to 0, 12, 24, 36, 48, and 60 hour timepoints. Black lines are present for image alignment.

**Supplementary Video 11: Time-lapse video of co-cultured WT (magenta) and PUM DKO (green) T-REx-293 cells.** Frames correspond to 30-minute intervals over a 55-hour period.

**Supplementary Video 12: Time-lapse video of co-cultured WT T-REx-293 cells (magenta) and WT 293T cells (green).** Frames correspond to 30-minute intervals over a 55-hour period.

## References


1 Wickens, M., Bernstein, D. S., Kimble, J. & Parker, R. A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet.* **18**, 150-157 (2002).

2 Lehmann, R. & Nüsslein-Volhard, C. Involvement of the pumilio gene in the transport of an abdominal signal in the Drosophila embryo. *Nature* **329**, 167, doi:10.1038/329167a0 (1987).

3 Nüsslein-Volhard, C., Frohnhöfer, H. G. & Lehmann, R. Determination of anteroposterior polarity in Drosophila. *Science* **238**, 1675-1681, doi:10.1126/science.3686007 (1987).

4 Zamore, P. D., Williamson, J. R. & Lehmann, R. The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**, 1421-1433 (1997).

5 Parisi, M. & Lin, H. The Drosophila pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis. *Genetics* **153**, 235-250 (1999).

6 Crittenden, S. L. *et al.* A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans. *Nature* **417**, 660-663, doi:10.1038/nature754 (2002).

7 Lublin, A. L. & Evans, T. C. The RNA-binding proteins PUF-5, PUF-6, and PUF-7 reveal multiple systems for maternal mRNA regulation during C. elegans oogenesis. *Dev. Biol.* **303**, 635-649, doi:10.1016/j.ydbio.2006.12.004 (2007).

8 Carreira-Rosario, A. *et al.* Repression of Pumilio Protein Expression by Rbfox1 Promotes Germ Cell Differentiation. *Dev. Cell* **36**, 562-571, doi:10.1016/j.devcel.2016.02.010 (2016).

9 Vessey, J. P. *et al.* Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. *J. Neurosci.* **26**, 6496-6508, doi:10.1523/jneurosci.0649-06.2006 (2006).

10 Sternburg, E. L., Estep, J. A., Nguyen, D. K., Li, Y. & Karginov, F. V. Antagonistic and cooperative AGO2-PUM interactions in regulating mRNAs. *Sci. Rep.* **8**, 15316, doi:10.1038/s41598-018-33596-4 (2018).

11 Bohn, J. A. *et al.* Identification of diverse target RNAs that are functionally regulated by human Pumilio proteins. *Nucleic Acids Res.* **46**, 362-386, doi:10.1093/nar/gkx1120 (2018).

12 Galgano, A. *et al.* Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. *PLoS One* **3**, e3164, doi:10.1371/journal.pone.0003164 (2008).

13 Zamore, P. D., Williamson, J. R. & Lehmann, R. The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**, 1421-1433 (1997).

14 Lu, G. & Hall, T. M. T. Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure* **19**, 361-367, doi:10.1016/j.str.2010.12.019 (2011).

15 Wang, X., McLachlan, J., Zamore, P. D. & Hall, T. M. T. Modular recognition of RNA by a human pumilio-homology domain. *Cell* **110**, 501-512, doi:10.1016/S0092-8674(02)00873-5 (2002).
16 Van Etten, J. *et al.* Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *J. Biol. Chem.* **287**, 36370-36383, doi:10.1074/jbc.M112.373522 (2012).
17 Sajek, M. *et al.* PUM1 and PUM2 exhibit different modes of regulation for SIAH1 that involve cooperativity with NANOS paralogues. *Cell. Mol. Life Sci.*, doi:10.1007/s00018-018-2926-5 (2018).
18 Jaruzelska, J. *et al.* Conservation of a Pumilio-Nanos complex from Drosophila germ plasm to human germ cells. *Dev. Genes Evol.* **213**, 120-126, doi:10.1007/s00427-003-0303-2 (2003).
19 Lin, K. *et al.* Essential requirement of mammalian PUMILIO family in embryonic development. *Mol. Biol. Cell*, mbcE18060369, doi:10.1091/mbc.E18-06-0369 (2018).
20 Zhang, M. *et al.* Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. *Genes Dev.*, doi:10.1101/gad.298752.117 (2017).
21 Mak, W., Xia, J., Cheng, E.-C., Lowther, K. & Lin, H. A role of Pumilio 1 in mammalian oocyte maturation and maternal phase of embryogenesis. *Cell Biosci.* **8**, 54, doi:10.1186/s13578-018-0251-1 (2018).
22 Mak, W., Fang, C., Holden, T., Dratver, M. B. & Lin, H. An Important Role of Pumilio 1 in Regulating the Development of the Mammalian Female Germline. *Biol. Reprod.* **94**, 134, doi:10.1095/biolreprod.115.137497 (2016).
23 Chen, D. *et al.* Pumilio 1 suppresses multiple activators of p53 to safeguard spermatogenesis. *Curr. Biol.* **22**, 420-425, doi:10.1016/j.cub.2012.01.039 (2012).
24 Zhang, M. *et al.* Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. *Genes Dev.* **31**, 1354-1369, doi:10.1101/gad.298752.117 (2017).
25 Zahr, S. K. *et al.* A Translational Repression Complex in Developing Mammalian Neural Stem Cells that Regulates Neuronal Specification. *Neuron* **97**, 520-537.e526, doi:10.1016/j.neuron.2017.12.045 (2018).
26 Gennarino, V. A. *et al.* A Mild PUM1 Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures. *Cell* **172**, 924-936.e911, doi:10.1016/j.cell.2018.02.006 (2018).
27 Gennarino, V. A. *et al.* Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels. *Cell* **160**, 1087-1098, doi:10.1016/j.cell.2015.02.012 (2015).
28 Liu, Y., Qu, L., Liu, Y., Roizman, B. & Zhou, G. G. PUM1 is a biphasic negative regulator of innate immunity genes by suppressing LGP2. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6902-E6911, doi:10.1073/pnas.1708713114 (2017).
29 Naudin, C. *et al.* PUMILIO/FOXP1 signaling drives expansion of hematopoietic stem/progenitor and leukemia cells. *Blood* **129**, 2493-2506, doi:10.1182/blood-2016-10-747436 (2017).

30 Miles, W. O., Tschop, K., Herr, A., Ji, J. Y. & Dyson, N. J. Pumilio facilitates miRNA regulation of the E2F3 oncogene. *Genes Dev.* **26**, 356-368, doi:10.1101/gad.182568.111 (2012).
31 Kedde, M. *et al.* A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol.* **12**, 1014-1020, doi:10.1038/ncb2105 (2010).
32 Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **164**, 69-80, doi:10.1016/j.cell.2015.12.017 (2016).
33 Kopp, F. *et al.* PUMILIO hyperactivity drives premature aging of Norad-deficient mice. *bioRxiv*, doi:10.1101/432112 (2018).
34 Goldstrohm, A. C., Hall, T. M. T. & McKenney, K. M. Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins. *Trends Genet.* **34**, 972-990, doi:10.1016/j.tig.2018.09.006 (2018).
35 Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **164**, 69-80, doi:10.1016/j.cell.2015.12.017 (2016).
36 Karperien, A., Ahammer, H. & Jelinek, H. F. Quantitating the subtleties of microglial morphology with fractal analysis. *Front. Cell. Neurosci.* **7**, 3, doi:10.3389/fncel.2013.00003 (2013).
37 Karperien, A. L. & Jelinek, H. F. Fractal, multifractal, and lacunarity analysis of microglia in tissue engineering. *Front Bioeng Biotechnol* **3**, 51, doi:10.3389/fbioe.2015.00051 (2015).
38 Smith, T. G., Lange, G. D. & Marks, W. B. Fractal methods and results in cellular morphology — dimensions, lacunarity and multifractals. *J. Neurosci. Methods* **69**, 123-136, doi:https://doi.org/10.1016/S0165-0270(96)00080-5 (1996).
39 Zahedi, A. *et al.* Deep Analysis of Mitochondria and Cell Health Using Machine Learning. *Sci. Rep.* **8**, 16354, doi:10.1038/s41598-018-34455-y (2018).
40 Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129-141, doi:10.1016/j.cell.2010.03.009 (2010).
41 Morris, A. R., Mukherjee, N. & Keene, J. D. Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets. *Mol. Cell. Biol.* **28**, 4093-4103, doi:10.1128/MCB.00155-08 (2008).
42 Jarmoskaite, I. *et al.* A quantitative and predictive model for RNA binding by human Pumilio proteins. *bioRxiv*, doi:10.1101/403006 (2018).
43 Vaidyanathan, P. P., AlSadhan, I., Merriman, D. K., Al-Hashimi, H. M. & Herschlag, D. Pseudouridine and N6-methyladenosine modifications weaken PUF protein/RNA interactions. *RNA* **23**, 611-618, doi:10.1261/rna.060053.116 (2017).
44 Friend, K. *et al.* A conserved PUF-Ago-eEF1A complex attenuates translation elongation. *Nat. Struct. Mol. Biol.* **19**, 176-183, doi:10.1038/nsmb.2214 (2012).
45 Wang, M. T. *et al.* K-Ras Promotes Tumorigenicity through Suppression of Non-canonical Wnt Signaling. *Cell* **163**, 1237-1251, doi:10.1016/j.cell.2015.10.041 (2015).

46 Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4**, 68-75 (2008).
47 Boylan, K. L. *et al.* The expression of Nectin-4 on the surface of ovarian cancer cells alters their ability to adhere, migrate, aggregate, and proliferate. *Oncotarget* **8**, 9717-9738, doi:10.18632/oncotarget.14206 (2017).
48 Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. Emerging Biological Principles of Metastasis. *Cell* **168**, 670-691, doi:10.1016/j.cell.2016.11.037 (2017).
49 Duff, D. & Long, A. Roles for RACK1 in cancer cell migration and invasion. *Cell. Signal.* **35**, 250-255, doi:10.1016/j.cellsig.2017.03.005 (2017).
50 Ritz, C. & Streibig, J. C. Nonlinear Regression with R Introduction. *Use R*, 1-+ (2008).
51 Mirman, D. *Growth curve analysis and visualization using R*. (CRC Press/Taylor & Francis Group, 2014).
52 Backman, T. W. H. & Girke, T. J. B. B. systemPipeR: NGS workflow and report generation environment. **17**, 388, doi:10.1186/s12859-016-1241-0 (2016).

# CHAPTER 4

# Conclusions

## Main conclusions and observations

This work showed extensive interaction between PUM and AGO, both at a site level and transcriptome-wide. Most AGO2-bound transcripts were occupied by one or both PUM proteins, and changes in PUM protein level lead to changes in AGO2 binding strength on co-occupied 3′ UTRs. AGO's reciprocal ability to affect PUM binding was also observed in the CLIP data. The identification of antagonistic co-regulation of two individual mRNAs supports a model where binding of PUM near the AGO2 sites leads to attenuation of miRNA-guided repression. This shows that Pumilio, a normally repressive factor, can take on a stabilizing role in a context-dependent manner, and underscores the flexibility of combinatorial interaction outcomes. This observation is also consistent with other studies that have suggested a stabilizing role for the Pumilio proteins[1-4], as well as with the RNA-sequencing data as part of this work.

In addition, this work has phenotypically characterized a role of PUM in cell adhesion and migration. PUM DKO cells predominantly remained attached once contact is made, whereas WT cells formed transient interactions and were able to detach and continue individual movement. This attachment defect can be rescued upon reintroduction of PUM1 or PUM2, as well as with addition of supplemental extracellular matrix to the growth

substrate. Defects in collective cell migration, as well as shifts in actin morphology, are also tied to loss of PUM. It is important to note that addition of extracellular matrix can only rescue the clumping phenotype and does not rescue the migration rate defect. This suggests that the two phenotypes are separable and may be rooted in misregulation of different factors.

There was a significant amount of overlap in the populations of transcripts bound by both PUM proteins and, in agreement with their extensive homology, binding sites of both paralogs showed enrichment of a nearly identical motif consistent with the known UGUAnAUA recognition sequence. More interestingly, differences in binding preference, RBP interactions, and downstream regulation between PUM1 and PUM2 were observed. PUM1 bound a substantially larger population of transcripts than PUM2, a result that is seen across datasets[5,6]. Overlap of individual PUM1 and PUM2 binding sites was significantly less common than co-occupancy anywhere on the same 3′ UTRs, with paralog-specific loci still enriched in the binding motif. Differential motif discovery between PUM1 and PUM2 sites showed an enrichment of miRNA seeds near PUM2 sites and had a higher degree of overlap with AGO2 sites, suggesting a greater extent of interaction between PUM2 and the miRNA machinery. In addition, PUM2 showed evidence of both cooperative and antagonistic interactions across transcripts, while PUM1 only showed evidence of antagonistic interactions. PUM1 displays an increased ability to rescue growth and adhesion phenotypes. This is supported by the results of the RNA-sequencing data, where more genes were differentially expressed upon addition of PUM1 that PUM2.

Differences between the two paralogs may be explained by differences in the N-terminal domains of each protein. PUM1 contains a longer N-terminal domain which may direct distinct behaviors, possibly through interactions with other proteins. This extended N-terminal domain may also mediate a higher binding affinity for PUM1.

## Broader implications

This work, alongside other studies, highlights the important role PUM proteins play in cellular function and development. Because of PUMs widespread regulatory effects, misregulation of PUM activity often leads to disease. Cancer stands out as one major disease which can potentially be tied to PUMs role in regulating cell division, adhesion, and motility. Cancer cells are known to increase malignancy through decreasing contacts with neighboring cells and in some cases upregulating motility genes[7,8]. Loss of PUM leads to a decrease in cell growth, an increase in cell adhesion, and a decrease in cell motility, effects associated with less aggressive cancers. These effects may explain why high expression of PUM proteins is associated with a lower chance of survival for patients with liver cancer. However, increased expression of PUM1 in renal cancers is associated with an increased chance of survival, a trend not explained by observations in this work. Variability in initial cell type and the underlying molecular mechanisms between cancers no doubt play a role in this inconsistency. In addition, PUM proteins have been shown to negatively regulate both tumor suppressor and proto-oncogenes, which can lead to different effects on cell function.

Candidate pathways have emerged that may underlie the phenotypes we observe. The Wnt signaling pathways regulated many aspects of cell-cell adhesion and migration across cell types. Frizzled-8 (FZD8), a non-canonical Wnt protein receptor, is increased in PUM DKO cells and may lead to an increase in downstream signaling. Independent expression analysis of PUM knockdown cells has also shown an enrichment in genes connected to Wnt signaling[2]. This pathway is capable of influencing activity of the family of Rho GTPases (mainly RhoA, Rac1, and Cdc42), and may explain the shifts in actin morphology observed. The mTOR pathway is important in proper regulation of cell growth and proliferation. The target gene Ras-related GTP-binding protein D (RRAGD) is a component of the amino acid sensing branch of mTORC1 signaling. Disruptions in this pathway may interfere with the cells ability to sense available nutrient levels and may lead to changes in growth rate.

PUM proteins also play a substantial role in neurogenesis, and misregulation of PUMs is associated with defects in brain development and function. Both PUM1 and PUM2 single knockout mice have defects in brain development, as well as brain-specific PUM double knockout mice[9]. Loss of PUM1 has been shown to increase the expression of Ataxin1 (ATXN1), a defect that in both mice and humans causes spinocerebellar ataxia[10,11]. Interestingly, misregulation of the target gene very low-density lipoprotein receptor (VLDLR) leads to similar neurodevelopmental disorders, including cerebellar ataxia[12,13]. VLDLR regulates dendritic spine formation, and loss of VLDLR has been demonstrated to disrupt neuronal migration in mice[14,15]. PUM DKO cells tested in this study also have

decreased VLDLR levels accompanied by defects in migration, although it is yet to be determined if the two are connected.

## Future directions

The precise mechanism(s) of AGO-PUM antagonism have yet to be determined. Involvement of secondary structure rearrangements of RNA is a good starting hypothesis that has been previously implicated[16-18]. Alternatively, steric clashes may be responsible. Additional mutagenesis of the two sites within the VLDLR and RRAGD 3' UTRs which change the distance and/or sequence context between binding sites will provide a more precise picture of the mechanism of interaction between AGO and PUM. Additionally, predicted sites of antagonism and cooperation can be compared transcriptome-wide to determine whether differences in the distance between AGO and PUM sites affect interaction behavior.

Candidates for transmitting PUM effects from the RNA-sequencing data have yet to be tested. For candidates that increase in expression upon PUM loss, siRNAs will be used to knockdown expression levels in PUM DKO cells. For candidates that decrease in expression, expression plasmids with the coding sequence of interest will be stably transfected into PUM DKO cells. If one of the candidates is responsible for the observed clumping and/or migration defects, then restoration of wildtype morphology and migration rate are expected. It is possible that multiple genes contribute to the observed phenotype, so a full phenotypic rescue may not be observed for each candidate.

Based on the results, defects in extracellular matrix synthesis and/or secretion are also likely. Mass spectrometry can be performed on collected extracellular matrix and other secreted proteins for both WT and PUM DKO cells in order to determine what components of the extracellular environment are disrupted. This can be compared to the RNA-sequencing data for further validation. Restoration of these factors to the extracellular environment would then restore normal morphology.

Finally, PUM proteins can be knocked out in the neuroblastoma cell line, N2A, in order to better assess the role of PUM proteins in the migration and cytoskeletal coordination within neurons. Based on the RNA-sequencing data in T-Rex-293 cells, it is hypothesized that PUM may play a role proper neurite outgrowth. Comparing the quantity, quality, and cytoskeletal behavior of neurites between WT and PUM DKO N2A cells upon differentiation can shed light on this.

## References


1 Kaye, J. A., Rose, N. C., Goldsworthy, B., Goga, A. & L'Etoile, N. D. A 3'UTR pumilio-binding element directs translational activation in olfactory sensory neurons. *Neuron* **61**, 57-70, doi:10.1016/j.neuron.2008.11.012 (2009).

2 Bohn, J. A. *et al.* Identification of diverse target RNAs that are functionally regulated by human Pumilio proteins. *Nucleic Acids Res.* **46**, 362-386, doi:10.1093/nar/gkx1120 (2018).

3 Naudin, C. *et al.* PUMILIO/FOXP1 signaling drives expansion of hematopoietic stem/progenitor and leukemia cells. *Blood* **129**, 2493-2506, doi:10.1182/blood-2016-10-747436 (2017).

4 Lee, C. D. & Tu, B. P. Glucose-Regulated Phosphorylation of the PUF Protein Puf3 Regulates the Translational Fate of Its Bound mRNAs and Association with RNA Granules. *Cell Rep.* **11**, 1638-1650, doi:10.1016/j.celrep.2015.05.014 (2015).

5 Galgano, A. *et al.* Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. *PLoS One* **3**, e3164, doi:10.1371/journal.pone.0003164 (2008).

6 Zhang, M. *et al.* Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. *Genes Dev.*, doi:10.1101/gad.298752.117 (2017).

7 Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. Emerging Biological Principles of Metastasis. *Cell* **168**, 670-691, doi:10.1016/j.cell.2016.11.037 (2017).

8 Duff, D. & Long, A. Roles for RACK1 in cancer cell migration and invasion. *Cell. Signal.* **35**, 250-255, doi:10.1016/j.cellsig.2017.03.005 (2017).

9 Zhang, M. *et al.* Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. *Genes Dev.* **31**, 1354-1369, doi:10.1101/gad.298752.117 (2017).

10 Gennarino, V. A. *et al.* A Mild PUM1 Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures. *Cell* **172**, 924-936.e911, doi:10.1016/j.cell.2018.02.006 (2018).

11 Gennarino, V. A. *et al.* Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels. *Cell* **160**, 1087-1098, doi:10.1016/j.cell.2015.02.012 (2015).

12 Boycott, K. M. *et al.* Homozygous deletion of the very low density lipoprotein receptor gene causes autosomal recessive cerebellar hypoplasia with cerebral gyral simplification. *Am. J. Hum. Genet.* **77**, 477-483, doi:10.1086/444400 (2005).

13 Schlotawa, L. *et al.* Cerebellar ataxia, mental retardation and dysequilibrium syndrome 1 (CAMRQ1) caused by an unusual constellation of VLDLR mutation. *J. Neurol.* **260**, 1678-1680, doi:10.1007/s00415-013-6941-z (2013).

14 Trommsdorff, M. *et al.* Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. *Cell* **97**, 689-701 (1999).
15 DiBattista, A. M. *et al.* Very low density lipoprotein receptor regulates dendritic spine formation in a RasGRF1/CaMKII dependent manner. *Biochim. Biophys. Acta* **1853**, 904-917, doi:10.1016/j.bbamcr.2015.01.015 (2015).
16 HafezQorani, S. *et al.* Modeling the combined effect of RNA-binding proteins and microRNAs in post-transcriptional regulation. *Nucleic Acids Res.* **44**, e83, doi:10.1093/nar/gkw048 (2016).
17 Kedde, M. *et al.* A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol.* **12**, 1014-1020, doi:10.1038/ncb2105 (2010).
18 Leibovich, L., Mandel-Gutfreund, Y. & Yakhini, Z. A structural-based statistical approach suggests a cooperative activity of PUM1 and miR-410 in human 3'-untranslated regions. *Silence* **1**, 17, doi:10.1186/1758-907X-1-17 (2010).

# APPENDICES

## Appendix A: AGO2 and PUM CLIP alignment and annotation tables.

**Table A1: AGO2 CLIP alignment and annotation table for PUM1 knockdown experiments.**

| Library: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **IP** | **Ago2** | **Ago2** | **Ago2** | **Ago2** | **Ago2** | **Ago2** | **Ago2** | **Ago2** | **Ago2** |
| **Experiment** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** | **PUM1 KD** |
| **siRNA** | **gl3.1** | **PUM1-1** | **PUM1-3** | **gl3.1** | **PUM1-1** | **PUM1-3** | **gl3.1** | **PUM1-1** | **PUM1-3** |
| **rep** | **1** | **1** | **1** | **2** | **2** | **2** | **3** | **3** | **3** |
| **Total sequenced reads in library** | 6450307 | 6779255 | 4354969 | 23351989 | 15821209 | 18873363 | 43540795 | 40793064 | 57289529 |
| **reads >15nt after clipping adaptor** | 5223508 | 4502536 | 3817472 | 19870501 | 13141250 | 15920783 | 29363346 | 29060482 | 46194027 |
| **Unique CLIP fragment sequences after collapsing** | 240598 | 331219 | 168740 | 1836243 | 1190550 | 1209037 | 2527500 | 2187945 | 2945018 |
| **Uniquely mapped sequences** | 60526 | 45800 | 52072 | 248120 | 178338 | 189676 | 167080 | 164190 | 256743 |
| **Total readcounts for uniquely mappped sequences** | 105600 | 75729 | 79385 | 344693 | 251339 | 266624 | 294243 | 306190 | 455270 |
| **Annotation breakdown:** | | | | | | | | | |
| **unannotated** | 12349 | 11380 | 8126 | 70551 | 48495 | 48639 | 59531 | 62385 | 82892 |
| **exon_CDS** | 19784 | 12773 | 23912 | 34064 | 24859 | 36926 | 29447 | 27608 | 44715 |
| **exon_non_coding** | 1601 | 1326 | 1195 | 15598 | 10919 | 9125 | 10780 | 9979 | 20598 |
| **exon_utr3** | 20959 | 14269 | 17537 | 44746 | 29701 | 34747 | 44064 | 43216 | 57491 |
| **exon_utr5** | 1469 | 1088 | 1249 | 3605 | 2342 | 2795 | 4306 | 3475 | 6283 |
| **intron** | 9515 | 9044 | 7523 | 117031 | 89434 | 90183 | 85956 | 82903 | 156659 |
| **mature_miRNA** | 34736 | 21884 | 16686 | 25894 | 23236 | 21270 | 33960 | 50728 | 44868 |
| **pre-miRNA** | 1190 | 762 | 534 | 1352 | 803 | 777 | 2077 | 2564 | 2538 |
| **simple** | 152 | 105 | 95 | 980 | 696 | 715 | 906 | 889 | 1170 |
| **structural_RNA** | 1211 | 793 | 700 | 4602 | 3015 | 2603 | 3156 | 3051 | 5515 |
| **transposon** | 2634 | 2305 | 1828 | 26270 | 17839 | 18844 | 20060 | 19392 | 32541 |

**Table A2: AGO2 CLIP alignment and annotation table for PUM2 knockdown experiments.**

| Library: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| IP | Ago2 | Ago2 | Ago2 | Ago2 | Ago2 | Ago2 | Ago2 | Ago2 | Ago2 |
| Experim. | PUM2 KD | PUM2 KD | PUM2 KD | PUM2 KD | PUM2 KD | PUM2 KD | PUM2 KD | PUM2 KD | PUM2 KD |
| siRNA | gl3.1 | PUM2-1 | PUM2-2 | gl3.1 | PUM2-1 | PUM2-2 | gl3.1 | PUM2-1 | PUM2-2 |
| rep | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 |
| Total sequenced reads in library | 23059179 | 22029821 | 30483210 | 36505433 | 26816687 | 19525756 | 33299062 | 33103020 | 28894126 |
| reads >15nt after clipping adaptor | 21687107 | 19938015 | 28349232 | 21312802 | 16433597 | 10523144 | 26196233 | 28763365 | 20289876 |
| Unique CLIP fragment sequences after collapsing | 500260 | 456326 | 435560 | 2137361 | 1764780 | 1724592 | 984531 | 935526 | 835400 |
| Uniquely mapped sequences | 86758 | 69205 | 79413 | 128112 | 188052 | 175831 | 105012 | 103468 | 75729 |
| Total readcounts for uniquely mappped sequences | 190569 | 145522 | 200858 | 212149 | 282011 | 240928 | 200101 | 196668 | 153246 |
| Annotation breakdown: | | | | | | | | | |
| unannotated | 30410 | 30556 | 30390 | 61054 | 63842 | 58923 | 42408 | 45136 | 33689 |
| exon_CDS | 19239 | 11782 | 16565 | 21138 | 38383 | 26473 | 37820 | 29016 | 20529 |
| exon_non_coding | 3431 | 2463 | 2565 | 12550 | 11967 | 16078 | 4315 | 4843 | 2966 |
| exon_utr3 | 35998 | 18320 | 32300 | 13577 | 22779 | 15538 | 32758 | 26805 | 20665 |
| exon_utr5 | 1159 | 728 | 998 | 3526 | 3150 | 4025 | 2156 | 1634 | 1024 |
| intron | 32226 | 26392 | 30551 | 61877 | 101666 | 83811 | 39034 | 43624 | 28522 |
| mature_miRNA | 56512 | 46364 | 77158 | 14901 | 13916 | 10223 | 28754 | 31488 | 36106 |
| pre-miRNA | 2418 | 1260 | 2969 | 1114 | 742 | 903 | 1023 | 1203 | 1010 |
| simple | 303 | 201 | 164 | 1349 | 911 | 1071 | 596 | 621 | 350 |
| structural_RNA | 1599 | 1164 | 1629 | 3137 | 4102 | 3909 | 1920 | 1852 | 1237 |
| transposon | 7274 | 6292 | 5569 | 17926 | 20553 | 19974 | 9317 | 10446 | 7148 |

**Table A2 (continued): AGO2 CLIP alignment and annotation table for PUM2 knockdown experiments.**

| Library: | | | |
|---|---|---|---|
| IP | Ago2 | Ago2 | Ago2 |
| Experim. | PUM2 KD | PUM2 KD | PUM2 KD |
| siRNA | gl3.1 | PUM 2-1 | PUM 2-2 |
| rep | 4 | 4 | 4 |
| Total sequenced reads in library | 25393247 | 32053233 | 27133614 |
| reads >15nt after clipping adaptor | 19183244 | 25599167 | 22113010 |
| Unique CLIP fragment sequences after collapsing | 1817801 | 1831604 | 1637400 |
| Uniquely mapped sequences | 121197 | 155719 | 194350 |
| Total readcounts for uniquely mappped sequences | 192523 | 248459 | 293496 |
| Annotation breakdown: | | | |
| unannotated | 38059 | 48223 | 64415 |
| exon_CDS | 17051 | 22580 | 29313 |
| exon_non_coding | 8328 | 10520 | 11031 |
| exon_utr3 | 23510 | 31483 | 43048 |
| exon_utr5 | 2645 | 3170 | 3704 |
| intron | 65638 | 83034 | 90842 |
| mature_miRNA | 19688 | 26655 | 25361 |
| pre-miRNA | 1042 | 1603 | 1531 |
| simple | 626 | 722 | 697 |
| structural_RNA | 2072 | 2893 | 2555 |
| transposon | 13864 | 17576 | 20999 |

**Table A3: PUM CLIP alignment and annotation table.**

| Library: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **IP** | **PUM1** | **PUM1** | **PUM1** | **PUM1** | **PUM1** | **PUM1** | **PUM2** | **PUM2** | **PUM2** |
| **Experiment** | **miRNA def.** | **miRNA def.** | **miRNA def.** | **miRNA def.** | **miRNA def.** | **miRNA def.** | **miRNA def.** | **miRNA def.** | **miRNA def.** |
| **Cell type** | **WT** | **NoDice 2-20** | **NoDice 4-25** | **WT** | **NoDice 2-20** | **NoDice 4-25** | **WT** | **NoDice 2-20** | **NoDice 4-25** |
| **Rep** | **1** | **1** | **1** | **2** | **2** | **2** | **1** | **1** | **1** |
| **Total sequenced reads in library** | 31306486 | 31561629 | 31648795 | 26888379 | 17220018 | 10245946 | 13546456 | 12099217 | 16129570 |
| **reads >15nt after clipping adaptor** | 22453938 | 22101993 | 28928873 | 23752053 | 15763986 | 9412494 | 12084215 | 11187328 | 14781806 |
| **Unique CLIP fragment sequences after collapsing** | 1289647 | 1228283 | 1877270 | 803859 | 710662 | 500859 | 514191 | 338041 | 274544 |
| **Uniquely mapped sequences** | 104302 | 118757 | 663821 | 170145 | 154683 | 139443 | 112499 | 77387 | 57770 |
| **Total readcounts for uniquely mappped sequences** | 156929 | 175410 | 939264 | 236274 | 226459 | 204813 | 208279 | 132007 | 129915 |
| **Annotation breakdown:** | | | | | | | | | |
| **unannotated** | 27481 | 38150 | 116099 | 36759 | 31320 | 29842 | 32081 | 21005 | 20023 |
| **exon_CDS** | 14845 | 16381 | 85773 | 22035 | 23632 | 19322 | 21848 | 14568 | 11425 |
| **exon_non_coding** | 6095 | 6360 | 36987 | 9210 | 10385 | 10455 | 5143 | 3471 | 2447 |
| **exon_utr3** | 22975 | 22150 | 79622 | 62728 | 49208 | 33864 | 32937 | 24207 | 16094 |
| **exon_utr5** | 1814 | 1866 | 7400 | 3048 | 4129 | 2278 | 2338 | 1025 | 881 |
| **intron** | 50638 | 56777 | 518476 | 70505 | 77531 | 80940 | 58072 | 37278 | 26591 |
| **mature_miRNA** | 18633 | 18692 | 7866 | 13139 | 11096 | 9093 | 40651 | 20833 | 42417 |
| **pre-miRNA** | 860 | 748 | 375 | 758 | 627 | 457 | 1475 | 634 | 2740 |
| **simple** | 514 | 653 | 1596 | 857 | 447 | 615 | 346 | 159 | 140 |
| **structural_RNA** | 2425 | 2339 | 7405 | 2962 | 2952 | 2612 | 2058 | 1018 | 1756 |
| **transposon** | 10649 | 11294 | 77665 | 14273 | 15132 | 15335 | 11330 | 7809 | 5401 |

**Table A3: PUM CLIP alignment and annotation table.**

## Appendix B: Sequence information for siRNAs used for PUM1 and PUM2 knockdown, guide RNAs used for generation of PUM double knockout cells.

**Table B1: Knockdown siRNA sequences**

| Name | Sequence |
|---|---|
| GL3.1 sense | CUUACGCUGAGUACUUCGAUU |
| GL3.1 antisense | UCGAAGUACUCAGCGUAAGUU |
| dsiPum1_1_guide | UAAACCAUUGGUCUUGUCCAU |
| dsiPum1_1_pass | GGACAAGACCAAUGGUUUAUU |
| dsiPum1_3_guide | UUCUCUUCAAGACACAUGCAA |
| dsiPum1_3_pass | GCAUGUGUCUUGAAGAGAAUU |
| dsiPum2_1_guide | UAUGAAUCUAGAACCAUGCUG |
| dsiPum2_1_pass | GCAUGGUUCUAGAUUCAUAUU |
| dsiPum2_2_guide | UUUGAGCACAUGACCAUCCAG |
| dsiPum2_2_pass | GGAUGGUCAUGUGCUCAAAUU |

**Table B2: sgRNA, homology arm, and knockout screen sequences.**

| Name | Sequence |
|---|---|
| *single guide RNA sequences:* | |
| PUM1_4 ts | CACCGTGTCTCGCCATTGATCACCC |
| PUM1_15 ts | CACCGCATTCCATATCGCAAACGAG |
| PUM1_4 bs | AAACGGGTGATCAATGGCGAGACAC |
| PUM1_15 bs | AAACCTCGTTTGCGATATGGAATGC |
| *Homology Region Amplification Primers:* | |
| PUM1_4 RHA FP | GGGTCTCAGTCCACCCAGGAATATTCCTTTTGGTCCATCTTTG |
| PUM1_15 LHA FP | GGGTCTCAGGCCAGAGCAAGACTCTGTCTCAAAAAAACAAAACAAAAC |
| PUM1_4 RHA RP | GGGTCTCTCCACTGTGGAATTTTACTTTGTAAGCCTGGTCACT |
| PUM1_15 LHA RP | GGGTCTCTCACGGTTTGCGATATGGAATGTCTGATGTCATGC |

| *KO Screening/Validation Primers:* | |
|---|---|
| PUM1_4 Val FP | CAATTGCTGAGGAAAGGAGCTCTAAGACA |
| PUM1_4 Val RP | AGTTTCCTGTGTAATAGCAGTTGAAAACATAATTA GAGGA |
| resistance_Val FP | AAGACAATAGCAGGCATGCTGGG |

## Appendix C: Sequence information for candidate sites and luciferase assay constructs

**Table C1: AGO2 peak coordinates within candidates 3'UTRs.** Genomic locations are based on human genome assembly GRCh37/hg19.

| | | Coordinates of AGO2 Peak | | |
|---|---|---|---|---|
| **Name** | **Chromosome** | **Start** | **Stop** | **Strand** |
| KLHL15 | chrX | 24004784 | 24004833 | - |
| RGMA | chr15 | 93586694 | 93586737 | - |
| HNRNPA0 | chr5 | 137088356 | 137088396 | - |
| RRAGD | chr6 | 90077593 | 90077618 | - |
| CDKN1B-1 | chr12 | 12874234 | 12874297 | + |
| FNIP1 | chr5 | 130977422 | 130977452 | - |
| ACTG1 | chr17 | 79477161 | 79477242 | - |
| HNRNPA2B1-2 | chr7 | 26231541 | 26231593 | - |
| RPA2 | chr1 | 28218086 | 28218185 | - |
| KIAA0907 | chr1 | 155882913 | 155882961 | - |
| DAGLA | chr11 | 61514363 | 61514384 | + |
| ZC3H12C | chr11 | 110037535 | 110037582 | + |
| ATP6V1G1 | chr9 | 117360144 | 117360209 | + |
| ADD3 | chr10 | 111895149 | 111895244 | + |
| CDKN1B-2 | chr12 | 12875103 | 12875175 | + |
| HNRNPA2B1-3 | chr7 | 26231485 | 26231540 | - |
| PNRC1 | chr6 | 89794100 | 89794130 | + |
| VLDLR | chr9 | 2654231 | 2654268 | + |
| ZIC2 | chr13 | 100638546 | 100638612 | + |
| ZNF367 | chr9 | 99150401 | 99150450 | - |
| FOXO1 | chr13 | 41132471 | 41132496 | - |
| TOB1 | chr17 | 48939676 | 48939715 | - |
| SNAPC1 | chr14 | 62262149 | 62262173 | + |
| DUSP1 | chr5 | 172195296 | 172195359 | - |
| HNRNPA2B1-1 | chr7 | 26231728 | 26231817 | - |
| PMEPA1 | chr20 | 56226720 | 56226764 | - |

**Table C2: Monomer sequences for candidate 3'UTR sites.** Each sequence includes the entire AGO2 peak sites from CLIP-seq, entire overlapping PUM CLIP signals, and any PUM motifs within 200 nucleotides. Mutant AGO2 sequences have shuffled AGO2 peak sequence. Mutant PUM motif sequences have TGT to ACA mutations within the PUM motif. Mutant miRNA seed sequences have three nucleotide mutations within positions 2-7 positions of the miRNA seed site.

| Name | Sequence |
|---|---|
| **Wildtype sequences:** | |
| ACTG1 | TAGGACCCAGTTTCCTTTCTTAGCTGATGTCTTTGGCCAGAA CACCGTGGGCTGTTACTTGCTTTGAGTTGGAAGCGGTTTGCA TTTACGCCTGTAAATGTATTCATTCTTAATTTATGTAAGGTTT TTT |
| ADD3 | TGTATTACAATGTATGTAGAAATAGTAACCTGTGAACTATG CTTTTCCATAACTTTTTAAAAATATATATATCTAAATGAATG CAATGTGCATAAATATTTTTTAAACATAACAGTGAACTATTG CACCTTTTGCTAATGCCTCTATTTACTTGCTTTGGCATAAAG AATGAGCCAATGAACCTCTGTGTCCTGTGGAAAAATGTATA AATGTTATCTGA |
| ATP6V1G1 | ATTATATAATAGGTCCTTCCACTTTTTGGAGAGTAGCAAATC TAGCTTTTTTGTACAGACTTAGAAATTATCTAAAGATTTCAT CTTTTTACCTCATATTTCTTAGGAATTTAATGGTTATATGTTG TCTTTTTTTCCTATGTCTTTTGGCTCAAGCAACATGTATATCA GTGTTGACT |
| CDKN1B-1 | GTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATATAGTT TTTACCTTTTATGTAGCACATAAACTTTGGGGAAGGGAGGG CAGGGTGGGGCTGAGGAACTGACGTGGAGCGGGGTATGAA GA |
| CDKN1B-2 | AAAAACCATTTGAAGTGTACCTGTGTACATAACTCTGTAAA AACACTGAAAAATTATACTAACTTATTTATGTTAAAAGATTT TTTTTAATCTAGACAATATACAAGCCAAAGTGGCATGTTTTG TGCATTTGTAAATGCTGTGTTGGGTAGAATAGGTTTTCCCCT CTTTTGTTAAAT |
| DAGLA | AGACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTAATA TTGCTGTAAAAAGGAGAGACAAATTAATATAGCTTATTCTA TAAATATATCTGTATATAAAGGTTTCTGTATATTGTATAGAG CTGTGTATAAACTGGATGTAGAAGCACGCTGGCTGCC |
| DUSP1 | CTTCACAAATGTCATTGTCTACTCCTAGAAGAACCAAATACC TCAATTTTTGTTTTTGAGTACTGTACTATCCTGTAAATATATC TTAAGCAGGTTTGTTTT |
| FNIP1 | CTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTCTTG TTCTGTATACATTTCTCAAATGTACACTTGTATTATAATAAC CTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGTC A |

| | |
|---|---|
| FOXO1 | TTTGTTTATTTTGTTATTTGCAAATTTGTACAAACATTTAAATGGTTCTAATTTC |
| HNRNPA0 | TTCTATGAAATCTACTTGGATCCCATGCCTGAAATTTGGAAGCATATGTACAAAAATCATTTTTACGTTTTATTTTTAATAAATCATTGT |
| HNRNPA2B1-1 | CAGAGCAGATGCAGAGAGCCATTTTGTGAATGGATTGGATTATTTAATAACATTACCTTACTGTGGAGGAAGGATTGTAAAAAAAAAATGCCTTTGAGACAGTTTCTTAG |
| HNRNPA2B1-2 | TCTCAAAGTTTTGAAAAGCTATTAGCCAGGATCATGGTGTAATAAGACATAACGTTTTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATTGTGTAA |
| HNRNPA2B1-3 | GTTTTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATTGTGTAATTATAGACTTTTTATTTTAAATAAGTTAAGGAGTGGGTAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTTGTATAAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAATTCATCAGGATTAAATTATGTAGATAACTTAAGGGA |
| KIAA0907 | GTTTGAGATATTGAACTGTCATTTTTGCACATTTGAATACTTTGCAGGCTGGCTTTGTATAAACTTATCCTCTGGTTTCCTATATGTTGTAAATATTTAGACCATAATTTCATTATAAATAAATCTATAAATATTCTGCTTGTGGTT |
| KLHL15 | AATGTTACTGGTTTTATCTACTTGTTTATTTTGTACAAAATACCCAGCGACACTAGGGATGTAAGCCCTCAGTTTT |
| PMEPA1 | GTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCACTTAACTTTAAAGACATATTTGCACAAAACCTTTGTTTAAAGATCTGCAATATTATAT |
| PNRC1 | ATACAAACAGCTTGTATTATATTTTATATTTTGTAAATACTGTATACCATGTATTATGTGTATATTGTTCATACTTGAGAGGTATATTATAGTTTTGTTATGAAAGTATGTATTTTGCCCTGCCCACATTGCAGGTGTTTTGTATATATACAATGGATAAATTTTAAGTGTGTGCTAAGG |
| RGMA | CTGCGTCCACGTGTCTGCGACCTGTGTGGAGTGTCACCGCGTGTACATACTGTAAATTATTTATTAATGGCTAAATGCAA |
| RPA2 | ACTTTTTGACACACTTGCCATGACGTGTGTTTCTGTGAACATGAAGTTCTGCGGTAGTGCCTCCAGGGGCAGAGGAAAAGAAGAAGTGTTACTGCATTTTGTACAAAATAAATACAGTCATATGTTTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGTTT |
| RRAGD | CTTCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAAACTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCATTACACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCAGGAATTGTACATATTTTGCACTT |

| | |
|---|---|
| SNAPC1 | ATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTATT<br>TTGTATATTTATTCTATTTTGTATATTGTTAAGTGCAATAAA<br>GTTTTTGCCTTGCT |
| TOB1 | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAA<br>ATATTAATTTTGTACCTATATTGTGCAATACTTGAAAAAAAC<br>GGTATAAAAG |
| VLDLR | CTTGACCGTTTTTATATTACTTTTGTAAATATTCTTGTCCACA<br>TTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAACC<br>CTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGC<br>ACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCT<br>ATAGTGTATACCACCTGTACATACATTGTATAGGCCATCTGT<br>AAATATCCCAGAGAACAATCACTATTCTTAAGCACTTTGAA<br>AATATTTCTATGTAAATTATTGTAAACTTTTTCAATGGTTGG<br>GACAATGGCAATAGGACAAAACGGGTTACTAAGATG |
| ZC3H12C | ATTTCTTGATACTGCACTATAGAGAAATGGTGATGGAGGAG<br>TTGTAAATGGTAACTTAAAATTTTTGTAAGATATTGTATATT<br>TTCCATTTTCCTGAAGGTAGTTTTCTTGGGGGGGCCTGTTAT<br>ATTATTAAG |
| ZIC2 | TTATGAGGCAACCTGATTGTAAACTTCATGTAACTATAGACT<br>GGAAAAAATGAGCCGTGCCAAAGTCTCCCTTCTGTTTCTTCA<br>GCACATTGACCCATAGCACACACATACACACCACCACCAAC<br>AACGCTTGTGAATGTATTTTTCTGTTAGCTGGGTTTACATGT<br>GATGTTTTAGTGCTTTTGCAAGTTCAATTTGTTAGTTCCTGTA<br>TGAAAGATTGTGGGGGAAAAATAAACGTCGTGCCGTTAGCT<br>TTTTCCGTAATAACACCCTTCCTTCTGTAAATACCCGTTACC<br>ATATTTATCCATTTGTAATTAAATTATGGTATTAACT |
| ZNF367 | ACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTACAA<br>AAGTACTGCCCAGCATACTTAAAAAGTAGATCCTTGGGCAT<br>AAGCTAAGCACCTTATTTGCTTATCATAGGCTGCTATTCTGT<br>AGAAATTTATGAAGAATGTTATTGCCCCAGAATATGGGGTG<br>AGAGAGAACTGCACTTTTTTAATATGGAAATGAATTCATCG<br>TAAAGTTTAAAATATTTTGTAAATATGGACTGCACAGTACA<br>GGGTAGAAAACTACATATTG |
| hunchback | TTGTTGTCGAAAATTGTACATAAGCCAA |
| **mutant AGO2 Sequence:** | |
| ACTG1 | TAGGACCCAGGGCTGTTCTGAGCTTGGTAGTATGTTCATTTC<br>TTCGATCTCTAGCGTATAGCCGCTCTGCCCAGACTTGTGGTT<br>GAATGTTCTGTAAATGTATTCATTCTTAATTTATGTAAGGTT<br>TTTT |
| ADD3 | TGTATTACAATGTATGTAGAAATAGTAACCTGTGAACTATG<br>CTTTTCCATAACTTTTTAAAAATATATATATCTAAATGAATG<br>CAATGTGCATAAATATTTTTTAAACTTCCGCCCTCCCCGTCG<br>TAAATGAAGGTAACTATTGAAAATATTAACTGACGCAGTTT |

| | |
|---|---|
| | AGTTCAGTCAATTGATGTTAGATTCTATGCCAGTATAGTATA AATGTTATCTGA |
| ATP6V1G1 | ATTATATAATTAAACACGGAAATCAGGTTCTAGTCGTTTTCA CTACCTTGATTATCATGATAAGTAAGTTGTTTCAGATTTCAT CTTTTTACCTCATATTTCTTAGGAATTTAATGGTTATATGTTG TCTTTTTTTCCTATGTCTTTTGGCTCAAGCAACATGTATATCA GTGTTGACT |
| CDKN1B-1 | GTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATATAGTT TTTACCTTTTATAAGTAGGGTGAGGGGGTGAAAAGGGTCGG TGCGCGGTATGGCTGCAGGGAGCAGTGACAACGGTATGAAG A |
| CDKN1B-2 | AAAAACCATTTGAAGTGTACCTGTGTACATAACTCTGTAAA AACACTGAAAAATTATACTAACTTATTTATGTTAAAAGATTT TTTTTAATCTAGACTGAACTGTAAGTCTCTGGAGAGTGCGTG TTGAGAATGTGTTACTTTAATTCCAGATATGAGACTTCTCTT CATTTGTTAAAT |
| DAGLA | AGACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTAATG AGAATATCAGAATGCGGAATAAATTAATATAGCTTATTCTA TAAATATATCTGTATATAAAGGTTTCTGTATATTGTATAGAG CTGTGTATAAACTGGATGTAGAAGCACGCTGGCTGCC |
| DUSP1 | CTTCACAAATGACCGTTTTATTTCGCACTTAACTTGCTTGAA CAACCTTTGTGTTCTTAGAACAAATACTCTATGTAAATATAT CTTAAGCAGGTTTGTTTT |
| FNIP1 | CTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTCTTG TTCTGTATACATTTCTCAAATGTACACTTGTATTATAATAAC CTCATGGTTCGGATTACTGATCAACTAGGTAATACACATGTC A |
| FOXO1 | TTTGTTTATTTGATCTTTAAATGGTTTAACATTCAATTTAAAT GGTTCTAATTTC |
| HNRNPA0 | TTCTATGAAAAGGTCCATTTTTAAGGTAACTATCATCCCGTT GCTGAAAGCAAAAATCATTTTTACGTTTTATTTTTAATAAAT CATTGT |
| HNRNPA2B1-1 | CAGAGCAGATTAGGTGAAGGCTTATAAATAGTATCGGGTTG GGGACTAAGCGTAAAATCAGCAAAATAGAGTCATAATCCTA TGTCTTAGAAATTTTTGAGTTTCTTAG |
| HNRNPA2B1-2 | TCTCAAAGTTCCAATTAATAGTGCTCTACCTGGTAATAAAAT ATTGCGTAGAGTTAGTCATGTTAAAAAAATTTAAGTGCGTG TGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATAAAA TAATTCTAAAGGAAATTGTGTAA |
| HNRNPA2B1-3 | GTTTTTCCTTATGTATTTGAAGGGTATGAGAATACATTAGTG GGTTTATTTTTTACTAAAAAGACAATAAAATAATTCTAAAG GAAATTGTGTAATTATAGACTTTTTATTTTAAATAAGTTAAG GAGTGGGTAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTT |

| | |
|---|---|
| | ATATTTGTATAAGATAAATGCTGGTCAGATGTAAGTGTGTTG<br>TCTGCAATTCATCAGGATTAAATTATGTAGATAACTTAAGG<br>GA |
| KIAA0907 | GTTTGAGATATAGTGTTCGGATTATTGTCCTATCTTCGTGAT<br>ATTGATTTCGATCGCAATAAACTTATCCTCTGGTTTCCTATA<br>TGTTGTAAATATTTAGACCATAATTTCATTATAAATAAATCT<br>ATAAATATTCTGCTTGTGGTT |
| KLHL15 | AATGTTACTGGTTTTATTAAGATTTAGTACGTCCCAGTAAGT<br>CCGAATGTCCTTATCGACATGTAACCTCAGTTTT |
| PMEPA1 | GTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCACTT<br>AATGCTTATAGCTATACCTTATGAACCCGTTAACATTTGAAA<br>TAAAATATTATAT |
| PNRC1 | ATACAAACAGCTTGTATTATATTTTATATTTTGTAAATACTG<br>TATACCATGTATTCGATTGTTATTGTTAGGTATAGTGTACAA<br>TATTATAGTTTTGTTATGAAAGTATGTATTTTGCCCTGCCCA<br>CATTGCAGGTGTTTTGTATATATACAATGGATAAATTTTAAG<br>TGTGTGCTAAGG |
| RGMA | CTGCGTCCACAGGGGGGTGCGTTGTCACGTTTATGTTCTACG<br>GCAGCCCTACTAAATTATTTATTAATGGCTAAATGCAA |
| RPA2 | ACTTTTTGACGCCTATAGGTGCTGTTCGCCGTTTAGGTAGAT<br>CCATTCTAGGAGGATAATATGAAGAAAAATCTACGTCGCTG<br>AATCATCAATCGGATGGTGGGATGCAAAATACAGTCATATG<br>TTTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTC<br>GTTT |
| RRAGD | CTTCTCTTTTAATTAGTTGAACCAAGACATAGAATAAAAACT<br>TGTATATTAACAGTGATGTGAAATCCATTGTCATTTCATTAC<br>ACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCG<br>TGTGAGAACTGCCAGGAATTGTACATATTTTGCACTT |
| SNAPC1 | ATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTATT<br>TTGTATATTTATTCTATTTTGTACAGAAATTTTATGATGTAAT<br>TGTTTGCCTTGCT |
| TOB1 | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAA<br>ATAATATTCTTACAGGAAATCGAACTAATGTATTAATTTCAT<br>GGTATAAAAG |
| VLDLR | CTTGACCGTTTTTATATTACTTTTGTAAATATTCTTGTCCACA<br>TTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAACC<br>CTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGC<br>ACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCT<br>ATAGTGTATACCACCTGTACATACATTGTATAGGCCATCTGT<br>AAATATCCCTTAAAGTGTCGAACCTCAAATATTAATACAAA<br>CGTATTTCTATGTAAATTATTGTAAACTTTTTCAATGGTTGG<br>GACAATGGCAATAGGACAAAACGGGTTACTAAGATG |

| | |
|---|---|
| ZC3H12C | ATTTCTTGATTTGGGCTAGGGTAATTAATGAATAGAACCATA<br>ATTTGATAGAGCGGGAAAATTTTTGTAAGATATTGTATATTT<br>TCCATTTTCCTGAAGGTAGTTTTCTTGGGGGGGCCTGTTATA<br>TTATTAAG |
| ZIC2 | TTATGAGGCAACCTGATTGTAAACTTCATGTCGATAGTTCTT<br>AACCCAACCACCAGCTATTGCGACTCTGACGATCCGATGAT<br>TACAGGCAATATTATAGCACACACATACACACCACCACCAA<br>CAACGCTTGTGAATGTATTTTTCTGTTAGCTGGGTTTACATG<br>TGATGTTTTAGTGCTTTTGCAAGTTCAATTTGTTAGTTCCTGT<br>ATGAAAGATTGTGGGGGAAAAATAAACGTCGTGCCGTTAGC<br>TTTTTCCGTAATAACACCCTTCCTTCTGTAAATACCCGTTAC<br>CATATTTATCCATTTGTAATTAAATTATGGTATTAACT |
| ZNF367 | ACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTACAA<br>AAGTACTGCCCAGCATACTTAAAAAGTAGATCCTTGGGCAA<br>CAAGACGTAGTGTTGACTCGCTAGATCTCATGTATATTTACA<br>TATTCTTTATGAAGAATGTTATTGCCCCAGAATATGGGGTGA<br>GAGAGAACTGCACTTTTTTAATATGGAAATGAATTCATCGT<br>AAAGTTTAAAATATTTTGTAAATATGGACTGCACAGTACAG<br>GGTAGAAAACTACATATTG |
| **mutant PUM motif Sequence:** | |
| FNIP1 | CTAAGTTACTTAGATGTTGGATAATAACATAGCTGTTTCTTG<br>TTCATAATACATTTCTCAAATGTACACTTGTATTATAATAAC<br>CTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGTC<br>A |
| RRAGD | CTTCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAAACT<br>ATAATATTAACAGTGATGTGAAATCCATTGTCATTTCATTAC<br>ACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCG<br>TGTGAGAACTGCCAGGAATATAACATATTTTGCACTT |
| TOB1 | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAA<br>ATATTAATTTATAACCTATATTGTGCAATACTTGAAAAAAAC<br>GGTATAAAAG |
| VLDLR | CTTGACCGTTTTTATATTACTTTATAAAATATTCTTGTCCACA<br>TTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAACC<br>CTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGC<br>ACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCT<br>ATAGTGTATACCACCATAACATACATTGTATAGGCCATCAT<br>AAAATATCCCAGAGAACAATCACTATTCTTAAGCACTTTGA<br>AAATATTTCTAATAAAATTATTGTAAACTTTTTCAATGGTTG<br>GGACAATGGCAATAGGACAAAACGGGTTACTAAGATG |
| hunchback | TTGTTGTCGAAAATACAACATAAGCCAA |
| **mutant miRNA seed Sequence:** | |
| RRAGD | CTTCTCTTTTATAAATAAAGTAAGGTGTTTGAAGCAAAAACT<br>TGTATATTAACAGTGATGTGAAATCCATTGTCATTTCATTAC |

| | |
|---|---|
| | ACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCG TGTGAGAACTGCCAGGAATTGTACATATTTTGCACTT |
| TOB1 | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAA ATATTAATTTTGTACCTATATTGACGAATACTTGAAAAAAAC GGTATAAAAG |

**Figure C1: Diagram of monomer sequence assembly.**

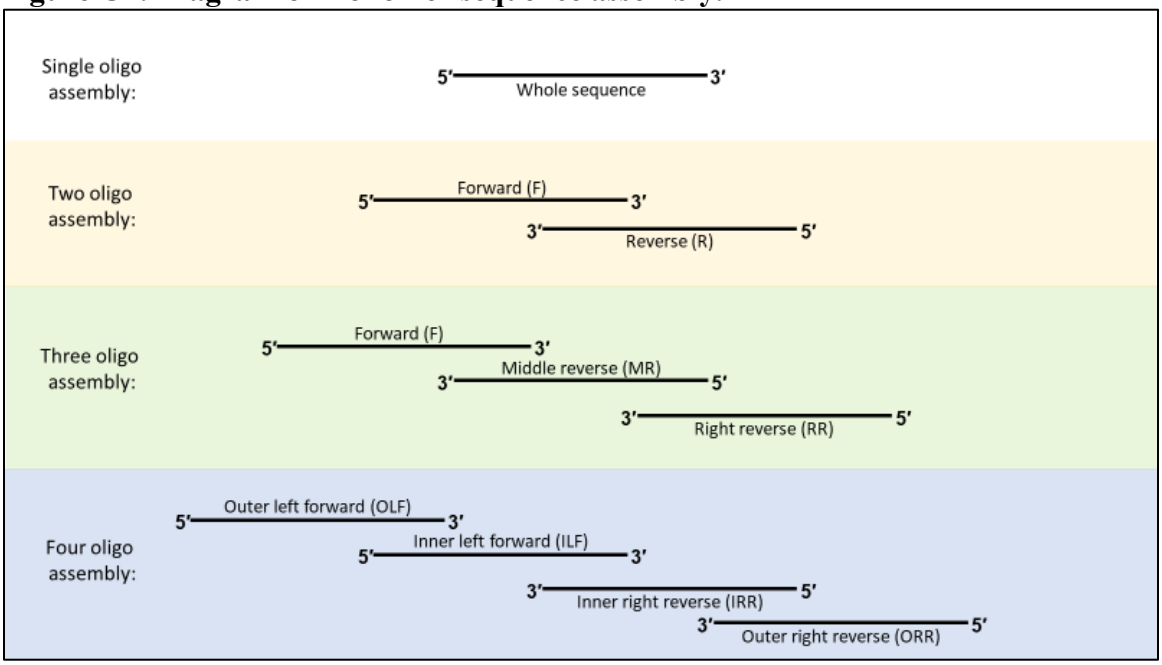



**Table C3: Oligos for monomer sequence assembly.** Monomer sequences were assembled from 1-4 initial oligos. Figure C1 denotes oligo nomenclature.

| **Name** | **Sequence** |
|---|---|
| ACTG1_F | TAGGACCCAGTTTCCTTTCTTAGCTGATGTCTTTGGCCAG AACACCGTGGGCTGTTACTTGCTTTGAGTT |
| ACTG1_R | AAAAAACCTTACATAAATTAAGAATGAATACATTTACAG GCGTAAATGCAAACCGCTTCCAACTCAAAGC |
| mACTG1_F | TAGGACCCAGGGCTGTTCTGAGCTTGGTAGTATGTTCATT TCTTCGATCTCTAGCGTATAGCCGCTCTGC |
| mACTG1_R | AAAAAACCTTACATAAATTAAGAATGAATACATTTACAG AACATTCAACCACAAGTCTGGGCAGAGCGGC |
| ADD3_LF | TGTATTACAATGTATGTAGAAATAGTAACCTGTGAACTAT GCTTTTCCATAACTTTTTAAAAATATATATATCTAAATG |

| | |
|---|---|
| ADD3_MR | TAAATAGAGGCATTAGCAAAAGGTGCAATAGTTCACTGTTATGTTTAAAAAATATTTATGCACATTGCATTCATTTAGATA |
| ADD3_RR | TCAGATAACATTTATACATTTTTCCACAGGACACAGAGGTTCATTGGCTCATTCTTTATGCCAAAGCAAGTAAATAGAGG |
| mADD3_MR | TATTTTCAATAGTTACCTTCATTTACGACGGGGAGGGCGGAAGTTTAAAAAATATTTATGCACATTGCATTCATTTAGATA |
| mADD3_RR | TCAGATAACATTTATACTATACTGGCATAGAATCTAACATCAATTGACTGAACTAAACTGCGTCAGTTAATATTTTCAAT |
| ATP6V1G1_LF | ATTATATAATAGGTCCTTCCACTTTTTGGAGAGTAGCAAATCTAGCTTTTTTGTACAGACTTAGA |
| ATP6V1G1_MR | TATAACCATTAAATTCCTAAGAAATATGAGGTAAAAAGATGAAATCTTTAGATAATTTCTAAGTCTG |
| ATP6V1G1_RR | AGTCAACACTGATATACATGTTGCTTGAGCCAAAAGACATAGGAAAAAAAGACAACATATAACCATT |
| mATP6V1G1_LF | ATTATATAATTAAACACGGAAATCAGGTTCTAGTCGTTTTCACTACCTTGATTATCATGATAAGT |
| mATP6V1G1_MR | TATAACCATTAAATTCCTAAGAAATATGAGGTAAAAAGATGAAATCTGAAACAACTTACTTATCATG |
| CDKN1B-1_F | GTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATATAGTTTTTACCTTTTATGTAGCACATAAAC |
| CDKN1B-1_R | TCTTCATACCCCGCTCCACGTCAGTTCCTCAGCCCCACCCTGCCCTCCCTTCCCCAAAGTTTATGTGC |
| mCDKN1B-1_F | GTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATATAGTTTTTACCTTTTATAAGTAGGGTGAGG |
| mCDKN1B-1_R | TCTTCATACCGTTGTCACTGCTCCCTGCAGCCATACCGCGCACCGACCCTTTTCACCCCCTCACCCTA |
| CDKN1B-2_LF | AAAAACCATTTGAAGTGTACCTGTGTACATAACTCTGTAAAAACACTGAAAAATTATACTAACTT |
| CDKN1B-2_MR | AACATGCCACTTTGGCTTGTATATTGTCTAGATTAAAAAAAATCTTTTAACATAAATAAGTTAGTAT |
| CDKN1B-2_RR | ATTTAACAAAAGAGGGGAAAACCTATTCTACCCAACACAGCATTTACAAATGCACAAAACATGCCAC |
| mCDKN1B-2_MR | GCACTCTCCAGAGACTTACAGTTCAGTCTAGATTAAAAAAAATCTTTTAACATAAATAAGTTAGTAT |
| mCDKN1B-2_RR | ATTTAACAAATGAAGAGAAGTCTCATATCTGGAATTAAAGTAACACATTCTCAACACGCACTCTCCA |
| DAGLA_F | AGACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTAATATTGCTGTAAAAAGGAGAGACAAATTAATATAGCTTATTCTATAA |

| | |
|---|---|
| DAGLA_R | GGCAGCCAGCGTGCTTCTACATCCAGTTTATACACAGCTC TATACAATATACAGAAACCTTTATATACAGATATATTTAT AGAATA |
| mDAGLA_F | AGACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTAA TGAGAATATCAGAATGCGGAATAAATTAATATAGCTTAT TCTATAA |
| DUSP1_F | CTTCACAAATGTCATTGTCTACTCCTAGAAGAACCAAATA CCTCAATTTTTGTTTT |
| DUSP1_R | AAAACAAACCTGCTTAAGATATATTTACAGGATAGTACA GTACTCAAAAACAAAAA |
| mDUSP1_F | CTTCACAAATGACCGTTTTATTTCGCACTTAACTTGCTTG AACAACCTTTGTGTTC |
| mDUSP1_R | AAAACAAACCTGCTTAAGATATATTTACATAGAGTATTTG TTCTAAGAACACAAAG |
| FNIP1_F | CTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTCT TGTTCTGTATACATTTCTCAAATGTACAC |
| FNIP1_R | TGACATGTGTATTTATTGCACAAATATCCCCTAGAACTGG GAGGTTATTATAATACAAGTGTACATTT |
| mFNIP1_R | TGACATGTGTATTACCTAGTTGATCAGTAATCCGAACCAT GAGGTTATTATAATACAAGTGTACATTT |
| mPUM_FNIP1_F | CTAAGTTACTTAGATGTTGGATAATAACATAGCTGTTTCT TGTTCATAATACATTTCTCAAATGTACAC |
| FOXO1_whole | TTTGTTTATTTTGTTATTTGCAAATTTGTACAAACATTTAA ATGGTTCTAATTTC |
| mFOXO1_whole | TTTGTTTATTTGATCTTTAAATGGTTTAACATTCAATTTAA ATGGTTCTAATTTC |
| HNRNPA0_whole | TTCTATGAAATCTACTTGGATCCCATGCCTGAAATTTGGA AGCATATGTACAAAAATCATTTTTACGTTTTATTTTTAAT AAATCATTGT |
| mHNRNPA0_whole | TTCTATGAAAAGGTCCATTTTTAAGGTAACTATCATCCCG TTGCTGAAAGCAAAAATCATTTTTACGTTTTATTTTTAAT AAATCATTGT |
| HNRNPA2B1-1_F | CAGAGCAGATGCAGAGAGCCATTTTGTGAATGGATTGGA TTATTTAATAACATTACCTTA |
| HNRNPA2B1-1_R | CTAAGAAACTGTCTCAAAGGCATTTTTTTTTACAATCCTT CCTCCACAGTAAGGTAATG |
| mHNRNPA2B1-1_F | CAGAGCAGATTAGGTGAAGGCTTATAAATAGTATCGGGT TGGGGACTAAGCGTAAAATCA |
| mHNRNPA2B1-1_R | CTAAGAAACTCAAAAATTTCTAAGACATAGGATTATGAC TCTATTTTGCTGATTTTACG |
| HNRNPA2B1-2_F | TCTCAAAGTTTTGAAAAGCTATTAGCCAGGATCATGGTGT AATAAGACATAACGTTTTTCCTTTAAAAAAATTTAAGTG |

| | |
|---|---|
| HNRNPA2B1-2_R | TTACACAATTTCCTTTAGAATTATTTTATTAAATCATAAATGTACAACAGCTTCTTAACTCTACACACGCACTTAAATT |
| mHNRNPA2B1-2_F | TCTCAAAGTTCCAATTAATAGTGCTCTACCTGGTAATAAAATATTGCGTAGAGTTAGTCATGTTAAAAAAATTTAAGTG |
| HNRNPA2B1-3_LF | GTTTTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATT |
| HNRNPA2B1-3_MR | TATAACAACAGCTTTGCAACGGACCTTAATTATACTACCCACTCCTTAACTTATTTAAAATAAAAAGTCTATAATTACACAATTTCCTTT |
| HNRNPA2B1-3_RR | TCCCTTAAGTTATCTACATAATTTAATCCTGATGAATTGCAGACAACACACTTACATCTGACCAGCATTTATCTTATACAAATATAACAA |
| mHNRNPA2B1-3_LF | GTTTTTCCTTATGTATTTGAAGGGTATGAGAATACATTAGTGGGTTTATTTTTTACTAAAAAGACAATAAAATAATTCTAAAGGAAATT |
| KIAA0907_F | GTTTGAGATATTGAACTGTCATTTTTGCACATTTGAATACTTTGCAGGCTGGCTTTGTATAAACTTATCCTCTGGTTTC |
| KIAA0907_R | AACCACAAGCAGAATATTTATAGATTTATTTATAATGAAATTATGGTCTAAATATTTACAACATATAGGAAACCAGAG |
| mKIAA0907_F | GTTTGAGATATAGTGTTCGGATTATTGTCCTATCTTCGTGATATTGATTTCGATCGCAATAAACTTATCCTCTGGTTTC |
| KLHL15_whole | AATGTTACTGGTTTTATCTACTTGTTTATTTTGTACAAAATACCCAGCGACACTAGGGATGTAAGCCCTCAGTTTT |
| mKLHL15_whole | AATGTTACTGGTTTTATTAAGATTTAGTACGTCCCAGTAAGTCCGAATGTCCTTATCGACATGTAACCTCAGTTTT |
| PMEPA1_F | GTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCACTTAACTTTAAAGA |
| PMEPA1_R | ATATAATATTGCAGATCTTTAAACAAAGGTTTTGTGCAAATATGTCTTTAAAGT |
| mPMEPA1_F | GTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCACTTAATGCTTATAG |
| mPMEPA1_R | ATATAATATTTTATTTCAAATGTTAACGGGTTCATAAGGTATAGCTATAAGCAT |
| PNRC1_LF | ATACAAACAGCTTGTATTATATTTTATATTTTGTAAATACTGTATACCATGTATTATGTGTATAT |
| PNRC1_MR | GCAGGGCAAAATACATACTTTCATAACAAAACTATAATATACCTCTCAAGTATGAACAATATACACAT |
| PNRC1_RR | CCTTAGCACACACTTAAAATTTATCCATTGTATATATACAAAACACCTGCAATGTGGGCAGGGCAAA |
| mPNRC1_LF | ATACAAACAGCTTGTATTATATTTTATATTTTGTAAATACTGTATACCATGTATTCGATTGTTAT |

| | |
|---|---|
| mPNRC1_MR | GCAGGGCAAAATACATACTTTCATAACAAAACTATAATA TTGTACACTATACCTAACAATAACAATCG |
| RGMA_whole | CTGCGTCCACGTGTCTGCGACCTGTGTGGAGTGTCACCGC GTGTACATACTGTAAATTATTTATTAATGGCTAAATGCAA |
| mRGMA_whole | CTGCGTCCACAGTGGTGCGGTGGTTGTCACGTATGTCTAC GTGCAGCCCTACTAAATTATTTATTAATGGCTAAATGCAA |
| RPA2_F | ACTTTTTGACACACTTGCCATGACGTGTGTTTCTGTGAAC ATGAAGTTCTGCGGTAGTGCCTCCAGGGGCAGAGGAAAA GAAGAAGTGTT |
| RPA2_R | AAACGAGAATTTTTACAAGTTACTACAATAGAACTGTTTT ATTAAACATATGACTGTATTTATTTTGTACAAAATGCAGT AACACTTCTT |
| mRPA2_F | ACTTTTTGACGCCTATAGGTGCTGTTCGCCGTTTAGGTAG ATCCATTCTAGGAGGATAATATGAAGAAAAATCTACGTC GCTGAATCATC |
| mRPA2_R | AAACGAGAATTTTTACAAGTTACTACAATAGAACTGTTTT ATTAAACATATGACTGTATTTTGCATCCCACCATCCGATT GATGATTCAG |
| RRAGD_F | CTTCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAAA CTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCA TTACAC |
| RRAGD_R | AAGTGCAAAATATGTACAATTCCTGGCAGTTCTCACACG GGATTTTTTTGACTACAGACCATAAAAGTTTACATTTGTG TAATGAA |
| mRRAGD_F | CTTCTCTTTTAATTAGTTGAACCAAGACATAGAATAAAAA CTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCA TTACAC |
| mmiR_RRAGD_F | CTTCTCTTTTATAAATAAAGTAAGGTGTTTGAAGCAAAAA CTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCA TTACAC |
| mPUM_RRAGD_F | CTTCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAAA CTATAATATTAACAGTGATGTGAAATCCATTGTCATTTCA TTACAC |
| mPUM_RRAGD_R | AAGTGCAAAATATGTTATATTCCTGGCAGTTCTCACACGG GATTTTTTTGACTACAGACCATAAAAGTTTACATTTGTGT AATGAA |
| SNAPC1_F | ATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTA TTTTGTATATTTAT |
| SNAPC1_R | AGCAAGGCAAAAACTTTATTGCACTTAACAATATACAAA ATAGAATAAATATAC |
| mSNAPC1_R | AGCAAGGCAAACAATTACATCATAAAATTTCTGTACAAA ATAGAATAAATATAC |

| | |
|---|---|
| TOB1_F | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAAATATTAATTTT |
| TOB1_R | CTTTTATACCGTTTTTTTCAAGTATTGCACAATATAGGTACAAAATTAATA |
| mTOB1_F | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAAATAATATTCTT |
| mTOB1_R | CTTTTATACCATGAAATTAATACATTAGTTCGATTTCCTGTAAGAATATTA |
| mmiR_TOB1_R | CTTTTATACCGTTTTTTTCAAGTATTCGTCAATATAGGTACAAAATTAATA |
| mPUM_TOB1_F | AGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGTAAATATTAATTTA |
| mPUM_TOB1_R | CTTTTATACCGTTTTTTTCAAGTATTGCACAATATAGGTTATAAATTAATA |
| VLDLR_ILF | TAACCCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGCACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCTA |
| VLDLR_IRR | CAAAGTGCTTAAGAATAGTGATTGTTCTCTGGGATATTTACAGATGGCCTATACAATGTATGTACAGGTGGTATACACTATAGCACAAGT |
| VLDLR_OLF | CTTGACCGTTTTTATATTACTTTTGTAAATATTCTTGTCCACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAACCCTTGA |
| VLDLR_ORR | CATCTTAGTAACCCGTTTTGTCCTATTGCCATTGTCCCAACCATTGAAAAAGTTTACAATAATTTACATAGAAATATTTTCAAAGTGCTT |
| mVLDLR_IRR | TGTATTAATATTTGAGGTTCGACACTTTAAGGGATATTTACAGATGGCCTATACAATGTATGTACAGGTGGTATACACTATAGCACAAGT |
| mVLDLR_ORR | CATCTTAGTAACCCGTTTTGTCCTATTGCCATTGTCCCAACCATTGAAAAAGTTTACAATAATTTACATAGAAATACGTTTGTATTAATA |
| mPUM_VLDLR_IRR | CAAAGTGCTTAAGAATAGTGATTGTTCTCTGGGATATTTTATGATGGCCTATACAATGTATGTTATGGTGGTATACACTATAGCACAAGT |
| mPUM_VLDLR_OLF | CTTGACCGTTTTTATATTACTTTATAAAATATTCTTGTCCACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAACCCTTGA |
| mPUM_VLDLR_ORR | CATCTTAGTAACCCGTTTTGTCCTATTGCCATTGTCCCAACCATTGAAAAAGTTTACAATAATTTTATTAGAAATATTTTCAAAGTGCTT |
| ZC3H12C_F | ATTTCTTGATACTGCACTATAGAGAAATGGTGATGGAGGAGTTGTAAATGGTAACTTAAAATTTTTGTAAGA |

| | |
|---|---|
| ZC3H12C_R | CTTAATAATATAACAGGCCCCCCCAAGAAAACTACCTTC AGGAAAATGGAAAATATACAATATCTTACAAAA |
| mZC3H12C_F | ATTTCTTGATTTGGGCTAGGGTAATTAATGAATAGAACCA TAATTTGATAGAGCGGGAAAATTTTTGTAAGA |
| ZIC2_ILF | TTCAGCACATTGACCCATAGCACACACATACACACCACC ACCAACAACGCTTGTGAATGTATTTTTCTGTTAGCTGGGT TTACATGTGAT |
| ZIC2_IRR | GCTAACGGCACGACGTTTATTTTTCCCCCACAATCTTTCA TACAGGAACTAACAAATTGAACTTGCAAAAGCACTAAAA CATCACATGTA |
| ZIC2_OLF | TTATGAGGCAACCTGATTGTAAACTTCATGTAACTATAGA CTGGAAAAAATGAGCCGTGCCAAAGTCTCCCTTCTGTTTC TTCAGCACAT |
| ZIC2_ORR | AGTTAATACCATAATTTAATTACAAATGGATAAATATGGT AACGGGTATTTACAGAAGGAAGGGTGTTATTACGGAAAA AGCTAACGGCA |
| mZIC2_ILF | GATTACAGGCAATATTATAGCACACACATACACACCACC ACCAACAACGCTTGTGAATGTATTTTTCTGTTAGCTGGGT TTACATGTGAT |
| mZIC2_OLF | TTATGAGGCAACCTGATTGTAAACTTCATGTCGATAGTTC TTAACCCAACCACCAGCTATTGCGACTCTGACGATCCGAT GATTACAGGC |
| ZNF367_ILF | AGTAGATCCTTGGGCATAAGCTAAGCACCTTATTTGCTTA TCATAGGCTGCTATTCTGTAGAAATTTATGAAGA |
| ZNF367_IRR | TGAATTCATTTCCATATTAAAAAAGTGCAGTTCTCTCTCA CCCCATATTCTGGGGCAATAACATTCTTCATAAA |
| ZNF367_OLF | ACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTACA AAAGTACTGCCCAGCATACTTAAAAAGTAGATCCT |
| ZNF367_ORR | CAATATGTAGTTTTCTACCCTGTACTGTGCAGTCCATATT TACAAAATATTTTAAACTTTACGATGAATTCATT |
| mZNF367_ILF | AGTAGATCCTTGGGCAACAAGACGTAGTGTTGACTCGCT AGATCTCATGTATATTTACATATTCTTTATGAAGA |
| hunchback | TTGTTGTCGAAAATTGTACATAAGCCAA |
| mPUM_ hunchback | TTGTTGTCGAAAATACAACATAAGCCAA |

**Figure C2: Diagram of restriction enzyme site addition to monomer ends.**

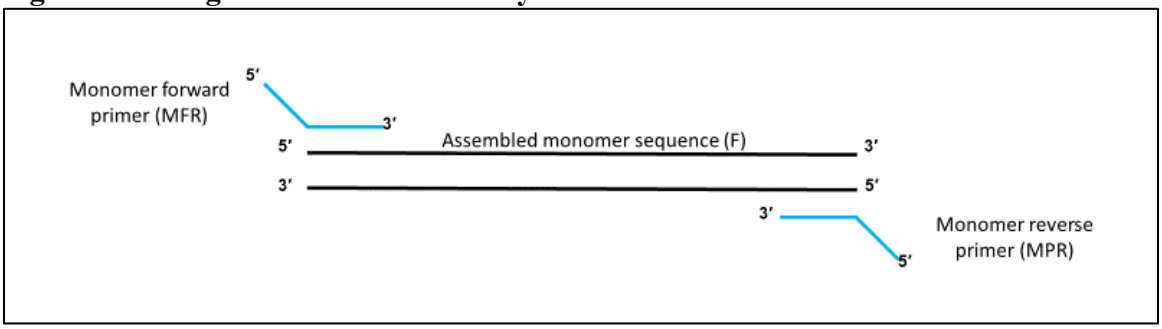



**Table C4: Primers for restriction enzyme (BsmBI) site addition to monomer ends.**
Figure C2 denotes primer nomenclature.

| Name | Sequence |
|---|---|
| ACTG1_M1FP | GCGTCTCTCACCTAGGACCCAG |
| ACTG1_M1RP | GCGTCTCTTGTTAAAAAACCTT |
| ACTG1_M2FP | GCGTCTCTAACATAGGACCCAG |
| ACTG1_M2RP | GCGTCTCTAGCCAAAAAACCTT |
| ACTG1_M3FP | GCGTCTCTGGCTTAGGACCCAG |
| ACTG1_M3RP | GCGTCTCTCTGAAAAAAACCTT |
| ACTG1_M4FP | GCGTCTCTTCAGTAGGACCCAG |
| ACTG1_M4RP | GCGTCTCTAAACAAAAAACCTT |
| ADD3_M1FP | GCGTCTCTCACCTGTATTACAA |
| ADD3_M1RP | GCGTCTCTTGTTTCAGATAACA |
| ADD3_M2FP | GCGTCTCTAACATGTATTACAA |
| ADD3_M2RP | GCGTCTCTAGCCTCAGATAACA |
| ADD3_M3FP | GCGTCTCTGGCTTGTATTACAA |
| ADD3_M3RP | GCGTCTCTCTGATCAGATAACA |
| ADD3_M4FP | GCGTCTCTTCAGTGTATTACAA |
| ADD3_M4RP | GCGTCTCTAAACTCAGATAACA |
| ATP6V1G1_M1FP | GCGTCTCTCACCATTATATAAT |
| ATP6V1G1_M1RP | GCGTCTCTTGTTAGTCAACACT |
| ATP6V1G1_M2FP | GCGTCTCTAACAATTATATAAT |
| ATP6V1G1_M2RP | GCGTCTCTAGCCAGTCAACACT |
| ATP6V1G1_M3FP | GCGTCTCTGGCTATTATATAAT |
| ATP6V1G1_M3RP | GCGTCTCTCTGAAGTCAACACT |
| ATP6V1G1_M4FP | GCGTCTCTTCAGATTATATAAT |
| ATP6V1G1_M4RP | GCGTCTCTAAACAGTCAACACT |

| | |
|---|---|
| CDKN1B-1_M1FP | GCGTCTCTCACCGTTTTTCCTT |
| CDKN1B-1_M1RP | GCGTCTCTTGTTTCTTCATACC |
| CDKN1B-1_M2FP | GCGTCTCTAACAGTTTTTCCTT |
| CDKN1B-1_M2RP | GCGTCTCTAGCCTCTTCATACC |
| CDKN1B-1_M3FP | GCGTCTCTGGCTGTTTTTCCTT |
| CDKN1B-1_M3RP | GCGTCTCTCTGATCTTCATACC |
| CDKN1B-1_M4FP | GCGTCTCTTCAGGTTTTTCCTT |
| CDKN1B-1_M4RP | GCGTCTCTAAACTCTTCATACC |
| CDKN1B-2_M1FP | GCGTCTCTCACCAAAAACCATT |
| CDKN1B-2_M1RP | GCGTCTCTTGTTATTTAACAAA |
| CDKN1B-2_M2FP | GCGTCTCTAACAAAAAACCATT |
| CDKN1B-2_M2RP | GCGTCTCTAGCCATTTAACAAA |
| CDKN1B-2_M3FP | GCGTCTCTGGCTAAAAACCATT |
| CDKN1B-2_M3RP | GCGTCTCTCTGAATTTAACAAA |
| CDKN1B-2_M4FP | GCGTCTCTTCAGAAAAACCATT |
| CDKN1B-2_M4RP | GCGTCTCTAAACATTTAACAAA |
| DAGLA_M1FP | GCGTCTCTCACCAGACTTTTTT |
| DAGLA_M1RP | GCGTCTCTTGTTGGCAGCCAGC |
| DAGLA_M2FP | GCGTCTCTAACAAGACTTTTTT |
| DAGLA_M2RP | GCGTCTCTAGCCGGCAGCCAGC |
| DAGLA_M3FP | GCGTCTCTGGCTAGACTTTTTT |
| DAGLA_M3RP | GCGTCTCTCTGAGGCAGCCAGC |
| DAGLA_M4FP | GCGTCTCTTCAGAGACTTTTTT |
| DAGLA_M4RP | GCGTCTCTAAACGGCAGCCAGC |
| DUSP1_M1FP | GCGTCTCTCACCCTTCACAAAT |
| DUSP1_M1RP | GCGTCTCTTGTTAAAACAAACC |
| DUSP1_M2FP | GCGTCTCTAACACTTCACAAAT |
| DUSP1_M2RP | GCGTCTCTAGCCAAAACAAACC |
| DUSP1_M3FP | GCGTCTCTGGCTCTTCACAAAT |
| DUSP1_M3RP | GCGTCTCTCTGAAAAACAAACC |
| DUSP1_M4FP | GCGTCTCTTCAGCTTCACAAAT |
| DUSP1_M4RP | GCGTCTCTAAACAAAACAAACC |
| FNIP1_M1FP | GCGTCTCTCACCCTAAGTTACT |
| FNIP1_M1RP | GCGTCTCTTGTTTGACATGTGT |
| FNIP1_M2FP | GCGTCTCTAACACTAAGTTACT |
| FNIP1_M2RP | GCGTCTCTAGCCTGACATGTGT |
| FNIP1_M3FP | GCGTCTCTGGCTCTAAGTTACT |

| | |
|---|---|
| FNIP1_M3RP | GCGTCTCTCTGATGACATGTGT |
| FNIP1_M4FP | GCGTCTCTTCAGCTAAGTTACT |
| FNIP1_M4RP | GCGTCTCTAAACTGACATGTGT |
| FOXO1_M1FP | GCGTCTCTCACCTTTGTTTATT |
| FOXO1_M1RP | GCGTCTCTTGTTGAAATTAGAA |
| FOXO1_M2FP | GCGTCTCTAACATTTGTTTATT |
| FOXO1_M2RP | GCGTCTCTAGCCGAAATTAGAA |
| FOXO1_M3FP | GCGTCTCTGGCTTTTGTTTATT |
| FOXO1_M3RP | GCGTCTCTCTGAGAAATTAGAA |
| FOXO1_M4FP | GCGTCTCTTCAGTTTGTTTATT |
| FOXO1_M4RP | GCGTCTCTAAACGAAATTAGAA |
| HNRNPA0_M1FP | GCGTCTCTCACCTTCTATGAAA |
| HNRNPA0_M1RP | GCGTCTCTTGTTACAATGATTT |
| HNRNPA0_M2FP | GCGTCTCTAACATTCTATGAAA |
| HNRNPA0_M2RP | GCGTCTCTAGCCACAATGATTT |
| HNRNPA0_M3FP | GCGTCTCTGGCTTTCTATGAAA |
| HNRNPA0_M3RP | GCGTCTCTCTGAACAATGATTT |
| HNRNPA0_M4FP | GCGTCTCTTCAGTTCTATGAAA |
| HNRNPA0_M4RP | GCGTCTCTAAACACAATGATTT |
| HNRNPA2B1-1_M1FP | GCGTCTCTCACCCAGAGCAGAT |
| HNRNPA2B1-1_M1RP | GCGTCTCTTGTTCTAAGAAACT |
| HNRNPA2B1-1_M2FP | GCGTCTCTAACACAGAGCAGAT |
| HNRNPA2B1-1_M2RP | GCGTCTCTAGCCCTAAGAAACT |
| HNRNPA2B1-1_M3FP | GCGTCTCTGGCTCAGAGCAGAT |
| HNRNPA2B1-1_M3RP | GCGTCTCTCTGACTAAGAAACT |
| HNRNPA2B1-1_M4FP | GCGTCTCTTCAGCAGAGCAGAT |
| HNRNPA2B1-1_M4RP | GCGTCTCTAAACCTAAGAAACT |
| HNRNPA2B1-2_M1FP | GCGTCTCTCACCTCTCAAAGTT |
| HNRNPA2B1-2_M1RP | GCGTCTCTTGTTTTACACAATT |
| HNRNPA2B1-2_M2FP | GCGTCTCTAACATCTCAAAGTT |
| HNRNPA2B1-2_M2RP | GCGTCTCTAGCCTTACACAATT |
| HNRNPA2B1-2_M3FP | GCGTCTCTGGCTTCTCAAAGTT |
| HNRNPA2B1-2_M3RP | GCGTCTCTCTGATTACACAATT |
| HNRNPA2B1-2_M4FP | GCGTCTCTTCAGTCTCAAAGTT |
| HNRNPA2B1-2_M4RP | GCGTCTCTAAACTTACACAATT |
| HNRNPA2B1-3_M1FP | GCGTCTCTCACCGTTTTTCCTT |
| HNRNPA2B1-3_M1RP | GCGTCTCTTGTTTCCCTTAAGT |

| | |
|---|---|
| HNRNPA2B1-3_M2FP | GCGTCTCTAACAGTTTTTCCTT |
| HNRNPA2B1-3_M2RP | GCGTCTCTAGCCTCCCTTAAGT |
| HNRNPA2B1-3_M3FP | GCGTCTCTGGCTGTTTTTCCTT |
| HNRNPA2B1-3_M3RP | GCGTCTCTCTGATCCCTTAAGT |
| HNRNPA2B1-3_M4FP | GCGTCTCTTCAGGTTTTTCCTT |
| HNRNPA2B1-3_M4RP | GCGTCTCTAAACTCCCTTAAGT |
| KIAA0907_M1FP | GCGTCTCTCACCGTTTGAGATA |
| KIAA0907_M1RP | GCGTCTCTTGTTAACCACAAGC |
| KIAA0907_M2FP | GCGTCTCTAACAGTTTGAGATA |
| KIAA0907_M2RP | GCGTCTCTAGCCAACCACAAGC |
| KIAA0907_M3FP | GCGTCTCTGGCTGTTTGAGATA |
| KIAA0907_M3RP | GCGTCTCTCTGAAACCACAAGC |
| KIAA0907_M4FP | GCGTCTCTTCAGGTTTGAGATA |
| KIAA0907_M4RP | GCGTCTCTAAACAACCACAAGC |
| KLHL15_M1FP | GCGTCTCTCACCAATGTTACTG |
| KLHL15_M1RP | GCGTCTCTTGTTAAAACTGAGG |
| KLHL15_M2FP | GCGTCTCTAACAAATGTTACTG |
| KLHL15_M2RP | GCGTCTCTAGCCAAAACTGAGG |
| KLHL15_M3FP | GCGTCTCTGGCTAATGTTACTG |
| KLHL15_M3RP | GCGTCTCTCTGAAAAACTGAGG |
| KLHL15_M4FP | GCGTCTCTTCAGAATGTTACTG |
| KLHL15_M4RP | GCGTCTCTAAACAAAACTGAGG |
| PMEPA1_M1FP | GCGTCTCTCACCGTGCGTGAAT |
| PMEPA1_M1RP | GCGTCTCTTGTTATATAATATT |
| PMEPA1_M2FP | GCGTCTCTAACAGTGCGTGAAT |
| PMEPA1_M2RP | GCGTCTCTAGCCATATAATATT |
| PMEPA1_M3FP | GCGTCTCTGGCTGTGCGTGAAT |
| PMEPA1_M3RP | GCGTCTCTCTGAATATAATATT |
| PMEPA1_M4FP | GCGTCTCTTCAGGTGCGTGAAT |
| PMEPA1_M4RP | GCGTCTCTAAACATATAATATT |
| PNRC1_M1FP | GCGTCTCTCACCATACAAACAG |
| PNRC1_M1RP | GCGTCTCTTGTTCCTTAGCACA |
| PNRC1_M2FP | GCGTCTCTAACAATACAAACAG |
| PNRC1_M2RP | GCGTCTCTAGCCCCTTAGCACA |
| PNRC1_M3FP | GCGTCTCTGGCTATACAAACAG |
| PNRC1_M3RP | GCGTCTCTCTGACCTTAGCACA |
| PNRC1_M4FP | GCGTCTCTTCAGATACAAACAG |

| | |
|---|---|
| PNRC1_M4RP | GCGTCTCTAAACCCTTAGCACA |
| RGMA_M1FP | GCGTCTCTCACCCTGCGTCCAC |
| RGMA_M1RP | GCGTCTCTTGTTTTGCATTTAG |
| RGMA_M2FP | GCGTCTCTAACACTGCGTCCAC |
| RGMA_M2RP | GCGTCTCTAGCCTTGCATTTAG |
| RGMA_M3FP | GCGTCTCTGGCTCTGCGTCCAC |
| RGMA_M3RP | GCGTCTCTCTGATTGCATTTAG |
| RGMA_M4FP | GCGTCTCTTCAGCTGCGTCCAC |
| RGMA_M4RP | GCGTCTCTAAACTTGCATTTAG |
| RPA2_M1FP | GCGTCTCTCACCACTTTTTGAC |
| RPA2_M1RP | GCGTCTCTTGTTAAACGAGAAT |
| RPA2_M2FP | GCGTCTCTAACAACTTTTTGAC |
| RPA2_M2RP | GCGTCTCTAGCCAAACGAGAAT |
| RPA2_M3FP | GCGTCTCTGGCTACTTTTTGAC |
| RPA2_M3RP | GCGTCTCTCTGAAAACGAGAAT |
| RPA2_M4FP | GCGTCTCTTCAGACTTTTTGAC |
| RPA2_M4RP | GCGTCTCTAAACAAACGAGAAT |
| RRAGD_M1FP | GCGTCTCTCACCCTTCTCTTTT |
| RRAGD_M1RP | GCGTCTCTTGTTAAGTGCAAAA |
| RRAGD_M2FP | GCGTCTCTAACACTTCTCTTTT |
| RRAGD_M2RP | GCGTCTCTAGCCAAGTGCAAAA |
| RRAGD_M3FP | GCGTCTCTGGCTCTTCTCTTTT |
| RRAGD_M3RP | GCGTCTCTCTGAAAGTGCAAAA |
| RRAGD_M4FP | GCGTCTCTTCAGCTTCTCTTTT |
| RRAGD_M4RP | GCGTCTCTAAACAAGTGCAAAA |
| SNAPC1_M1FP | GCGTCTCTCACCATAACTATTT |
| SNAPC1_M1RP | GCGTCTCTTGTTAGCAAGGCAA |
| SNAPC1_M2FP | GCGTCTCTAACAATAACTATTT |
| SNAPC1_M2RP | GCGTCTCTAGCCAGCAAGGCAA |
| SNAPC1_M3FP | GCGTCTCTGGCTATAACTATTT |
| SNAPC1_M3RP | GCGTCTCTCTGAAGCAAGGCAA |
| SNAPC1_M4FP | GCGTCTCTTCAGATAACTATTT |
| SNAPC1_M4RP | GCGTCTCTAAACAGCAAGGCAA |
| TOB1_M1FP | GCGTCTCTCACCAGATTTTTGC |
| TOB1_M1RP | GCGTCTCTTGTTCTTTTATACC |
| TOB1_M2FP | GCGTCTCTAACAAGATTTTTGC |
| TOB1_M2RP | GCGTCTCTAGCCCTTTTATACC |

| | |
|---|---|
| TOB1_M3FP | GCGTCTCTGGCTAGATTTTTGC |
| TOB1_M3RP | GCGTCTCTCTGACTTTTATACC |
| TOB1_M4FP | GCGTCTCTTCAGAGATTTTTGC |
| TOB1_M4RP | GCGTCTCTAAACCTTTTATACC |
| VLDLR_M1FP | GCGTCTCTCACCCTTGACCGTT |
| VLDLR_M1RP | GCGTCTCTTGTTCATCTTAGTA |
| VLDLR_M2FP | GCGTCTCTAACACTTGACCGTT |
| VLDLR_M2RP | GCGTCTCTAGCCCATCTTAGTA |
| VLDLR_M3FP | GCGTCTCTGGCTCTTGACCGTT |
| VLDLR_M3RP | GCGTCTCTCTGACATCTTAGTA |
| VLDLR_M4FP | GCGTCTCTTCAGCTTGACCGTT |
| VLDLR_M4RP | GCGTCTCTAAACCATCTTAGTA |
| ZC3H12C_M1FP | GCGTCTCTCACCATTTCTTGAT |
| ZC3H12C_M1RP | GCGTCTCTTGTTCTTAATAATA |
| ZC3H12C_M2FP | GCGTCTCTAACAATTTCTTGAT |
| ZC3H12C_M2RP | GCGTCTCTAGCCCTTAATAATA |
| ZC3H12C_M3FP | GCGTCTCTGGCTATTTCTTGAT |
| ZC3H12C_M3RP | GCGTCTCTCTGACTTAATAATA |
| ZC3H12C_M4FP | GCGTCTCTTCAGATTTCTTGAT |
| ZC3H12C_M4RP | GCGTCTCTAAACCTTAATAATA |
| ZIC2_M1FP | GCGTCTCTCACCTTATGAGGCA |
| ZIC2_M1RP | GCGTCTCTTGTTAGTTAATACC |
| ZIC2_M2FP | GCGTCTCTAACATTATGAGGCA |
| ZIC2_M2RP | GCGTCTCTAGCCAGTTAATACC |
| ZIC2_M3FP | GCGTCTCTGGCTTTATGAGGCA |
| ZIC2_M3RP | GCGTCTCTCTGAAGTTAATACC |
| ZIC2_M4FP | GCGTCTCTTCAGTTATGAGGCA |
| ZIC2_M4RP | GCGTCTCTAAACAGTTAATACC |
| ZNF367_M1FP | GCGTCTCTCACCACTCCGACAG |
| ZNF367_M1RP | GCGTCTCTTGTTCAATATGTAG |
| ZNF367_M2FP | GCGTCTCTAACAACTCCGACAG |
| ZNF367_M2RP | GCGTCTCTAGCCCAATATGTAG |
| ZNF367_M3FP | GCGTCTCTGGCTACTCCGACAG |
| ZNF367_M3RP | GCGTCTCTCTGACAATATGTAG |
| ZNF367_M4FP | GCGTCTCTTCAGACTCCGACAG |
| ZNF367_M4RP | GCGTCTCTAAACCAATATGTAG |
| hunchback_M1_FP | GCGTCTCTCACCTTGTTGTCGA |

| | |
|---|---|
| hunchback_M1_RP | GCGTCTCTTGTTTTGGCTTATG |
| hunchback_M2_FP | GCGTCTCTAACATTGTTGTCGA |
| hunchback_M2_RP | GCGTCTCTAGCCTTGGCTTATG |
| hunchback_M3_FP | GCGTCTCTGGCTTTGTTGTCGA |
| hunchback_M3_RP | GCGTCTCTCTGATTGGCTTATG |
| hunchback_M4_FP | GCGTCTCTTCAGTTGTTGTCGA |
| hunchback_M4_RP | GCGTCTCTAAACTTGGCTTATG |

**Table C5: 4x Assembled Sequences.**

| Name | 4x Assembled Sequence |
|---|---|
| ACTG1 | CACCTAGGACCCAGTTTCCTTTCTTAGCTGATGTCTTTGGCCAG<br>AACACCGTGGGCTGTTACTTGCTTTGAGTTGGAAGCGGTTTGC<br>ATTTACGCCTGTAAATGTATTCATTCTTAATTTATGTAAGGTTT<br>TTTAACATAGGACCCAGTTTCCTTTCTTAGCTGATGTCTTTGGC<br>CAGAACACCGTGGGCTGTTACTTGCTTTGAGTTGGAAGCGGTT<br>TGCATTTACGCCTGTAAATGTATTCATTCTTAATTTATGTAAGG<br>TTTTTTGGCTTAGGACCCAGTTTCCTTTCTTAGCTGATGTCTTT<br>GGCCAGAACACCGTGGGCTGTTACTTGCTTTGAGTTGGAAGCG<br>GTTTGCATTTACGCCTGTAAATGTATTCATTCTTAATTTATGTA<br>AGGTTTTTTTCAGTAGGACCCAGTTTCCTTTCTTAGCTGATGTC<br>TTTGGCCAGAACACCGTGGGCTGTTACTTGCTTTGAGTTGGAA<br>GCGGTTTGCATTTACGCCTGTAAATGTATTCATTCTTAATTTAT<br>GTAAGGTTTTTTGTTT |
| mACTG1 | CACCTAGGACCCAGGGCTGTTCTGAGCTTGGTAGTATGTTCAT<br>TTCTTCGATCTCTAGCGTATAGCCGCTCTGCCCAGACTTGTGGT<br>TGAATGTTCTGTAAATGTATTCATTCTTAATTTATGTAAGGTTT<br>TTTAACATAGGACCCAGGGCTGTTCTGAGCTTGGTAGTATGTT<br>CATTTCTTCGATCTCTAGCGTATAGCCGCTCTGCCCAGACTTGT<br>GGTTGAATGTTCTGTAAATGTATTCATTCTTAATTTATGTAAGG<br>TTTTTTGGCTTAGGACCCAGGGCTGTTCTGAGCTTGGTAGTATG<br>TTCATTTCTTCGATCTCTAGCGTATAGCCGCTCTGCCCAGACTT<br>GTGGTTGAATGTTCTGTAAATGTATTCATTCTTAATTTATGTAA<br>GGTTTTTTTCAGTAGGACCCAGGGCTGTTCTGAGCTTGGTAGT<br>ATGTTCATTTCTTCGATCTCTAGCGTATAGCCGCTCTGCCCAGA<br>CTTGTGGTTGAATGTTCTGTAAATGTATTCATTCTTAATTTATG<br>TAAGGTTTTTTGTTT |
| ADD3 | CACCTGTATTACAATGTATGTAGAAATAGTAACCTGTGAACTA<br>TGCTTTTCCATAACTTTTTAAAAATATATATATCTAAATGAATG<br>CAATGTGCATAAATATTTTTTAAACATAACAGTGAACTATTGC |

| | |
|---|---|
| | ACCTTTTGCTAATGCCTCTATTTACTTGCTTTGGCATAAAGAAT<br>GAGCCAATGAACCTCTGTGTCCTGTGGAAAAATGTATAAATGT<br>TATCTGAAACATGTATTACAATGTATGTAGAAATAGTAACCTG<br>TGAACTATGCTTTTCCATAACTTTTTAAAAATATATATATCTAA<br>ATGAATGCAATGTGCATAAATATTTTTTAAACATAACAGTGAA<br>CTATTGCACCTTTTGCTAATGCCTCTATTTACTTGCTTTGGCAT<br>AAAGAATGAGCCAATGAACCTCTGTGTCCTGTGGAAAAATGT<br>ATAAATGTTATCTGAGGCTTGTATTACAATGTATGTAGAAATA<br>GTAACCTGTGAACTATGCTTTTCCATAACTTTTTAAAAATATAT<br>ATATCTAAATGAATGCAATGTGCATAAATATTTTTTAAACATA<br>ACAGTGAACTATTGCACCTTTTGCTAATGCCTCTATTTACTTGC<br>TTTGGCATAAAGAATGAGCCAATGAACCTCTGTGTCCTGTGGA<br>AAAATGTATAAATGTTATCTGATCAGTGTATTACAATGTATGT<br>AGAAATAGTAACCTGTGAACTATGCTTTTCCATAACTTTTTAA<br>AAATATATATATCTAAATGAATGCAATGTGCATAAATATTTTT<br>TAAACATAACAGTGAACTATTGCACCTTTTGCTAATGCCTCTA<br>TTTACTTGCTTTGGCATAAAGAATGAGCCAATGAACCTCTGTG<br>TCCTGTGGAAAAATGTATAAATGTTATCTGAGTTT |
| mADD3 | CACCTGTATTACAATGTATGTAGAAATAGTAACCTGTGAACTA<br>TGCTTTTCCATAACTTTTTAAAAATATATATATCTAAATGAATG<br>CAATGTGCATAAATATTTTTTAAACTTCCGCCCTCCCCGTCGTA<br>AATGAAGGTAACTATTGAAAATATTAACTGACGCAGTTTAGTT<br>CAGTCAATTGATGTTAGATTCTATGCCAGTATAGTATAAATGT<br>TATCTGAAACATGTATTACAATGTATGTAGAAATAGTAACCTG<br>TGAACTATGCTTTTCCATAACTTTTTAAAAATATATATATCTAA<br>ATGAATGCAATGTGCATAAATATTTTTTAAACTTCCGCCCTCCC<br>CGTCGTAAATGAAGGTAACTATTGAAAATATTAACTGACGCAG<br>TTTAGTTCAGTCAATTGATGTTAGATTCTATGCCAGTATAGTAT<br>AAATGTTATCTGAGGCTTGTATTACAATGTATGTAGAAATAGT<br>AACCTGTGAACTATGCTTTTCCATAACTTTTTAAAAATATATAT<br>ATCTAAATGAATGCAATGTGCATAAATATTTTTTAAACTTCCG<br>CCCTCCCCGTCGTAAATGAAGGTAACTATTGAAAATATTAACT<br>GACGCAGTTTAGTTCAGTCAATTGATGTTAGATTCTATGCCAG<br>TATAGTATAAATGTTATCTGATCAGTGTATTACAATGTATGTA<br>GAAATAGTAACCTGTGAACTATGCTTTTCCATAACTTTTTAAA<br>AATATATATATCTAAATGAATGCAATGTGCATAAATATTTTTT<br>AAACTTCCGCCCTCCCCGTCGTAAATGAAGGTAACTATTGAAA<br>ATATTAACTGACGCAGTTTAGTTCAGTCAATTGATGTTAGATT<br>CTATGCCAGTATAGTATAAATGTTATCTGAGTTT |
| ATP6V1G1 | CACCATTATATAATAGGTCCTTCCACTTTTTGGAGAGTAGCAA<br>ATCTAGCTTTTTTGTACAGACTTAGAAATTATCTAAAGATTTCA<br>TCTTTTTACCTCATATTTCTTAGGAATTTAATGGTTATATGTTGT<br>CTTTTTTTCCTATGTCTTTTGGCTCAAGCAACATGTATATCAGT<br>GTTGACTAACAATTATATAATAGGTCCTTCCACTTTTTGGAGA |

| | |
|---|---|
| | GTAGCAAATCTAGCTTTTTTGTACAGACTTAGAAATTATCTAA<br>AGATTTCATCTTTTTACCTCATATTTCTTAGGAATTTAATGGTT<br>ATATGTTGTCTTTTTTTCCTATGTCTTTTGGCTCAAGCAACATG<br>TATATCAGTGTTGACTGGCTATTATATAATAGGTCCTTCCACTT<br>TTTGGAGAGTAGCAAATCTAGCTTTTTTGTACAGACTTAGAAA<br>TTATCTAAAGATTTCATCTTTTTACCTCATATTTCTTAGGAATT<br>TAATGGTTATATGTTGTCTTTTTTTCCTATGTCTTTTGGCTCAAG<br>CAACATGTATATCAGTGTTGACTTCAGATTATATAATAGGTCC<br>TTCCACTTTTTGGAGAGTAGCAAATCTAGCTTTTTTGTACAGAC<br>TTAGAAATTATCTAAAGATTTCATCTTTTTACCTCATATTTCTT<br>AGGAATTTAATGGTTATATGTTGTCTTTTTTTCCTATGTCTTTTG<br>GCTCAAGCAACATGTATATCAGTGTTGACTGTTT |
| mATP6V1G1 | CACCATTATATAATTAAACACGGAAATCAGGTTCTAGTCGTTT<br>TCACTACCTTGATTATCATGATAAGTAAGTTGTTTCAGATTTCA<br>TCTTTTTACCTCATATTTCTTAGGAATTTAATGGTTATATGTTGT<br>CTTTTTTTCCTATGTCTTTTGGCTCAAGCAACATGTATATCAGT<br>GTTGACTAACAATTATATAATTAAACACGGAAATCAGGTTCTA<br>GTCGTTTTCACTACCTTGATTATCATGATAAGTAAGTTGTTTCA<br>GATTTCATCTTTTTACCTCATATTTCTTAGGAATTTAATGGTTA<br>TATGTTGTCTTTTTTTCCTATGTCTTTTGGCTCAAGCAACATGT<br>ATATCAGTGTTGACTGGCTATTATATAATTAAACACGGAAATC<br>AGGTTCTAGTCGTTTTCACTACCTTGATTATCATGATAAGTAAG<br>TTGTTTCAGATTTCATCTTTTTACCTCATATTTCTTAGGAATTTA<br>ATGGTTATATGTTGTCTTTTTTTCCTATGTCTTTTGGCTCAAGC<br>AACATGTATATCAGTGTTGACTTCAGATTATATAATTAAACAC<br>GGAAATCAGGTTCTAGTCGTTTTCACTACCTTGATTATCATGAT<br>AAGTAAGTTGTTTCAGATTTCATCTTTTTACCTCATATTTCTTA<br>GGAATTTAATGGTTATATGTTGTCTTTTTTTCCTATGTCTTTTGG<br>CTCAAGCAACATGTATATCAGTGTTGACTGTTT |
| CDKN1B-1 | CACCGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATATA<br>GTTTTTACCTTTTATGTAGCACATAAACTTTGGGGAAGGGAGG<br>GCAGGGTGGGGCTGAGGAACTGACGTGGAGCGGGGTATGAAG<br>AAACAGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATAT<br>AGTTTTTACCTTTTATGTAGCACATAAACTTTGGGGAAGGGAG<br>GGCAGGGTGGGGCTGAGGAACTGACGTGGAGCGGGGTATGAA<br>GAGGCTGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATA<br>TAGTTTTTACCTTTTATGTAGCACATAAACTTTGGGGAAGGGA<br>GGGCAGGGTGGGGCTGAGGAACTGACGTGGAGCGGGGTATGA<br>AGATCAGGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTAT<br>ATAGTTTTTACCTTTTATGTAGCACATAAACTTTGGGGAAGGG<br>AGGGCAGGGTGGGGCTGAGGAACTGACGTGGAGCGGGGTATG<br>AAGAGTTT |
| mCDKN1B-1 | CACCGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATATA<br>GTTTTTACCTTTTATAAGTAGGGTGAGGGGGTGAAAAGGGTCG |

| | |
|---|---|
| | GTGCGCGGTATGGCTGCAGGGAGCAGTGACAACGGTATGAAG<br>AAACAGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATAT<br>AGTTTTTACCTTTTATAAGTAGGGTGAGGGGGTGAAAAGGGTC<br>GGTGCGCGGTATGGCTGCAGGGAGCAGTGACAACGGTATGAA<br>GAGGCTGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTATA<br>TAGTTTTTACCTTTTATAAGTAGGGTGAGGGGGTGAAAAGGGT<br>CGGTGCGCGGTATGGCTGCAGGGAGCAGTGACAACGGTATGA<br>AGATCAGGTTTTTCCTTATTTGCTTCATTGTACTACCTGTGTAT<br>ATAGTTTTTACCTTTTATAAGTAGGGTGAGGGGGTGAAAAGGG<br>TCGGTGCGCGGTATGGCTGCAGGGAGCAGTGACAACGGTATG<br>AAGAGTTT |
| CDKN1B-2 | CACCAAAAACCATTTGAAGTGTACCTGTGTACATAACTCTGTA<br>AAAACACTGAAAAATTATACTAACTTATTTATGTTAAAAGATT<br>TTTTTTAATCTAGACAATATACAAGCCAAAGTGGCATGTTTTG<br>TGCATTTGTAAATGCTGTGTTGGGTAGAATAGGTTTTCCCCTCT<br>TTTGTTAAATAACAAAAAACCATTTGAAGTGTACCTGTGTACA<br>TAACTCTGTAAAAACACTGAAAAATTATACTAACTTATTTATG<br>TTAAAAGATTTTTTTTAATCTAGACAATATACAAGCCAAAGTG<br>GCATGTTTTGTGCATTTGTAAATGCTGTGTTGGGTAGAATAGG<br>TTTTCCCCTCTTTTGTTAAATGGCTAAAAACCATTTGAAGTGTA<br>CCTGTGTACATAACTCTGTAAAAACACTGAAAAATTATACTAA<br>CTTATTTATGTTAAAAGATTTTTTTTAATCTAGACAATATACAA<br>GCCAAAGTGGCATGTTTTGTGCATTTGTAAATGCTGTGTTGGG<br>TAGAATAGGTTTTCCCCTCTTTTGTTAAATTCAGAAAAACCATT<br>TGAAGTGTACCTGTGTACATAACTCTGTAAAAACACTGAAAAA<br>TTATACTAACTTATTTATGTTAAAAGATTTTTTTTAATCTAGAC<br>AATATACAAGCCAAAGTGGCATGTTTTGTGCATTTGTAAATGC<br>TGTGTTGGGTAGAATAGGTTTTCCCCTCTTTTGTTAAATGTTT |
| mCDKN1B-2 | CACCAAAAACCATTTGAAGTGTACCTGTGTACATAACTCTGTA<br>AAAACACTGAAAAATTATACTAACTTATTTATGTTAAAAGATT<br>TTTTTTAATCTAGACTGAACTGTAAGTCTCTGGAGAGTGCGTG<br>TTGAGAATGTGTTACTTTAATTCCAGATATGAGACTTCTCTTCA<br>TTTGTTAAATAACAAAAAACCATTTGAAGTGTACCTGTGTACA<br>TAACTCTGTAAAAACACTGAAAAATTATACTAACTTATTTATG<br>TTAAAAGATTTTTTTTAATCTAGACTGAACTGTAAGTCTCTGGA<br>GAGTGCGTGTTGAGAATGTGTTACTTTAATTCCAGATATGAGA<br>CTTCTCTTCATTTGTTAAATGGCTAAAAACCATTTGAAGTGTAC<br>CTGTGTACATAACTCTGTAAAAACACTGAAAAATTATACTAAC<br>TTATTTATGTTAAAAGATTTTTTTTAATCTAGACTGAACTGTAA<br>GTCTCTGGAGAGTGCGTGTTGAGAATGTGTTACTTTAATTCCA<br>GATATGAGACTTCTCTTCATTTGTTAAATTCAGAAAAACCATTT<br>GAAGTGTACCTGTGTACATAACTCTGTAAAAACACTGAAAAAT<br>TATACTAACTTATTTATGTTAAAAGATTTTTTTTAATCTAGACT |

| | |
|---|---|
| | GAACTGTAAGTCTCTGGAGAGTGCGTGTTGAGAATGTGTTACT<br>TTAATTCCAGATATGAGACTTCTCTTCATTTGTTAAATGTTT |
| DAGLA | CACCAGACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTA<br>ATATTGCTGTAAAAAGGAGAGACAAATTAATATAGCTTATTCT<br>ATAAATATATCTGTATATAAAGGTTTCTGTATATTGTATAGAG<br>CTGTGTATAAACTGGATGTAGAAGCACGCTGGCTGCCAACAA<br>GACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTAATATTG<br>CTGTAAAAAGGAGAGACAAATTAATATAGCTTATTCTATAAAT<br>ATATCTGTATATAAAGGTTTCTGTATATTGTATAGAGCTGTGTA<br>TAAACTGGATGTAGAAGCACGCTGGCTGCCGGCTAGACTTTTT<br>TTGTACTTAATGTATGAAAGATCCAAACTAATATTGCTGTAAA<br>AAGGAGAGACAAATTAATATAGCTTATTCTATAAATATATCTG<br>TATATAAAGGTTTCTGTATATTGTATAGAGCTGTGTATAAACT<br>GGATGTAGAAGCACGCTGGCTGCCTCAGAGACTTTTTTTGTAC<br>TTAATGTATGAAAGATCCAAACTAATATTGCTGTAAAAAGGAG<br>AGACAAATTAATATAGCTTATTCTATAAATATATCTGTATATA<br>AAGGTTTCTGTATATTGTATAGAGCTGTGTATAAACTGGATGT<br>AGAAGCACGCTGGCTGCCGTTT |
| mDAGLA | CACCAGACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTA<br>ATGAGAATATCAGAATGCGGAATAAATTAATATAGCTTATTCT<br>ATAAATATATCTGTATATAAAGGTTTCTGTATATTGTATAGAG<br>CTGTGTATAAACTGGATGTAGAAGCACGCTGGCTGCCAACAA<br>GACTTTTTTTGTACTTAATGTATGAAAGATCCAAACTAATGAG<br>AATATCAGAATGCGGAATAAATTAATATAGCTTATTCTATAAA<br>TATATCTGTATATAAAGGTTTCTGTATATTGTATAGAGCTGTGT<br>ATAAACTGGATGTAGAAGCACGCTGGCTGCCGGCTAGACTTTT<br>TTTGTACTTAATGTATGAAAGATCCAAACTAATGAGAATATCA<br>GAATGCGGAATAAATTAATATAGCTTATTCTATAAATATATCT<br>GTATATAAAGGTTTCTGTATATTGTATAGAGCTGTGTATAAAC<br>TGGATGTAGAAGCACGCTGGCTGCCTCAGAGACTTTTTTTGTA<br>CTTAATGTATGAAAGATCCAAACTAATGAGAATATCAGAATGC<br>GGAATAAATTAATATAGCTTATTCTATAAATATATCTGTATAT<br>AAAGGTTTCTGTATATTGTATAGAGCTGTGTATAAACTGGATG<br>TAGAAGCACGCTGGCTGCCGTTT |
| DUSP1 | CACCCTTCACAAATGTCATTGTCTACTCCTAGAAGAACCAAAT<br>ACCTCAATTTTTGTTTTTGAGTACTGTACTATCCTGTAAATATA<br>TCTTAAGCAGGTTTGTTTTAACACTTCACAAATGTCATTGTCTA<br>CTCCTAGAAGAACCAAATACCTCAATTTTTGTTTTTGAGTACTG<br>TACTATCCTGTAAATATATCTTAAGCAGGTTTGTTTTGGCTCTT<br>CACAAATGTCATTGTCTACTCCTAGAAGAACCAAATACCTCAA<br>TTTTTGTTTTTGAGTACTGTACTATCCTGTAAATATATCTTAAG<br>CAGGTTTGTTTTTCAGCTTCACAAATGTCATTGTCTACTCCTAG<br>AAGAACCAAATACCTCAATTTTTGTTTTTGAGTACTGTACTATC<br>CTGTAAATATATCTTAAGCAGGTTTGTTTTGTTT |

| | |
|---|---|
| mDUSP1 | CACCCTTCACAAATGACCGTTTTATTTCGCACTTAACTTGCTTG<br>AACAACCTTTGTGTTCTTAGAACAAATACTCTATGTAAATATA<br>TCTTAAGCAGGTTTGTTTTAACACTTCACAAATGACCGTTTTAT<br>TTCGCACTTAACTTGCTTGAACAACCTTTGTGTTCTTAGAACAA<br>ATACTCTATGTAAATATATCTTAAGCAGGTTTGTTTTGGCTCTT<br>CACAAATGACCGTTTTATTTCGCACTTAACTTGCTTGAACAAC<br>CTTTGTGTTCTTAGAACAAATACTCTATGTAAATATATCTTAAG<br>CAGGTTTGTTTTTCAGCTTCACAAATGACCGTTTTATTTCGCAC<br>TTAACTTGCTTGAACAACCTTTGTGTTCTTAGAACAAATACTCT<br>ATGTAAATATATCTTAAGCAGGTTTGTTTTGTTT |
| FNIP1 | CACCCTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTCT<br>TGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAATAAC<br>CTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGTCA<br>AACACTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTC<br>TTGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAATAA<br>CCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGTC<br>AGGCTCTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTT<br>CTTGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAATA<br>ACCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGT<br>CATCAGCTAAGTTACTTAGATGTTGGATATGTACATAGCTGTT<br>TCTTGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAAT<br>AACCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATG<br>TCAGTTT |
| mFNIP1 | CACCCTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTCT<br>TGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAATAAC<br>CTCATGGTTCGGATTACTGATCAACTAGGTAATACACATGTCA<br>AACACTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTTC<br>TTGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAATAA<br>CCTCATGGTTCGGATTACTGATCAACTAGGTAATACACATGTC<br>AGGCTCTAAGTTACTTAGATGTTGGATATGTACATAGCTGTTT<br>CTTGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAATA<br>ACCTCATGGTTCGGATTACTGATCAACTAGGTAATACACATGT<br>CATCAGCTAAGTTACTTAGATGTTGGATATGTACATAGCTGTT<br>TCTTGTTCTGTATACATTTCTCAAATGTACACTTGTATTATAAT<br>AACCTCATGGTTCGGATTACTGATCAACTAGGTAATACACATG<br>TCAGTTT |
| mP_FNIP1 | CACCCTAAGTTACTTAGATGTTGGATAATAACATAGCTGTTTC<br>TTGTTCATAATACATTTCTCAAATGTACACTTGTATTATAATAA<br>CCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGTC<br>AAACACTAAGTTACTTAGATGTTGGATAATAACATAGCTGTTT<br>CTTGTTCATAATACATTTCTCAAATGTACACTTGTATTATAATA<br>ACCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATGT<br>CAGGCTCTAAGTTACTTAGATGTTGGATAATAACATAGCTGTT<br>TCTTGTTCATAATACATTTCTCAAATGTACACTTGTATTATAAT |

| | |
|---|---|
| | AACCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACATG<br>TCATCAGCTAAGTTACTTAGATGTTGGATAATAACATAGCTGT<br>TTCTTGTTCATAATACATTTCTCAAATGTACACTTGTATTATAA<br>TAACCTCCCAGTTCTAGGGGATATTTGTGCAATAAATACACAT<br>GTCAGTTT |
| FOXO1 | CACCTTTGTTTATTTTGTTATTTGCAAATTTGTACAAACATTTA<br>AATGGTTCTAATTTCAACATTTGTTTATTTTGTTATTTGCAAAT<br>TTGTACAAACATTTAAATGGTTCTAATTTCGGCTTTTGTTTATT<br>TTGTTATTTGCAAATTTGTACAAACATTTAAATGGTTCTAATTT<br>CTCAGTTTGTTTATTTTGTTATTTGCAAATTTGTACAAACATTT<br>AAATGGTTCTAATTTCGTTT |
| mFOXO1 | CACCTTTGTTTATTTGATCTTTAAATGGTTTAACATTCAATTTA<br>AATGGTTCTAATTTCAACATTTGTTTATTTGATCTTTAAATGGT<br>TTAACATTCAATTTAAATGGTTCTAATTTCGGCTTTTGTTTATTT<br>GATCTTTAAATGGTTTAACATTCAATTTAAATGGTTCTAATTTC<br>TCAGTTTGTTTATTTGATCTTTAAATGGTTTAACATTCAATTTA<br>AATGGTTCTAATTTCGTTT |
| HNRNPA<br>0 | CACCTTCTATGAAATCTACTTGGATCCCATGCCTGAAATTTGG<br>AAGCATATGTACAAAAATCATTTTTACGTTTTATTTTTAATAAA<br>TCATTGTAACATTCTATGAAATCTACTTGGATCCCATGCCTGA<br>AATTTGGAAGCATATGTACAAAAATCATTTTTACGTTTTATTTT<br>TAATAAATCATTGTGGCTTTCTATGAAATCTACTTGGATCCCAT<br>GCCTGAAATTTGGAAGCATATGTACAAAAATCATTTTTACGTT<br>TTATTTTTAATAAATCATTGTTCAGTTCTATGAAATCTACTTGG<br>ATCCCATGCCTGAAATTTGGAAGCATATGTACAAAAATCATTT<br>TTACGTTTTATTTTTAATAAATCATTGTGTTT |
| mHNRNP<br>A0 | CACCTTCTATGAAAAGGTCCATTTTTAAGGTAACTATCATCCC<br>GTTGCTGAAAGCAAAAATCATTTTTACGTTTTATTTTTAATAAA<br>TCATTGTAACATTCTATGAAAAGGTCCATTTTTAAGGTAACTA<br>TCATCCCGTTGCTGAAAGCAAAAATCATTTTTACGTTTTATTTT<br>TAATAAATCATTGTGGCTTTCTATGAAAAGGTCCATTTTTAAG<br>GTAACTATCATCCCGTTGCTGAAAGCAAAAATCATTTTTACGT<br>TTTATTTTTAATAAATCATTGTTCAGTTCTATGAAAAGGTCCAT<br>TTTTAAGGTAACTATCATCCCGTTGCTGAAAGCAAAAATCATT<br>TTTACGTTTTATTTTTAATAAATCATTGTGTTT |
| HNRNPA<br>2B1-1 | CACCCAGAGCAGATGCAGAGAGCCATTTTGTGAATGGATTGG<br>ATTATTTAATAACATTACCTTACTGTGGAGGAAGGATTGTAAA<br>AAAAAATGCCTTTGAGACAGTTTCTTAGAACACAGAGCAGAT<br>GCAGAGAGCCATTTTGTGAATGGATTGGATTATTTAATAACAT<br>TACCTTACTGTGGAGGAAGGATTGTAAAAAAAAATGCCTTTGA<br>GACAGTTTCTTAGGGCTCAGAGCAGATGCAGAGAGCCATTTTG<br>TGAATGGATTGGATTATTTAATAACATTACCTTACTGTGGAGG<br>AAGGATTGTAAAAAAAAATGCCTTTGAGACAGTTTCTTAGTCA<br>GCAGAGCAGATGCAGAGAGCCATTTTGTGAATGGATTGGATT |

| | |
|---|---|
| | ATTTAATAACATTACCTTACTGTGGAGGAAGGATTGTAAAAAA<br>AAATGCCTTTGAGACAGTTTCTTAGGTTT |
| mHNRNP<br>A2B1-1 | CACCCAGAGCAGATTAGGTGAAGGCTTATAAATAGTATCGGG<br>TTGGGGACTAAGCGTAAAATCAGCAAAATAGAGTCATAATCC<br>TATGTCTTAGAAATTTTTGAGTTTCTTAGAACACAGAGCAGAT<br>TAGGTGAAGGCTTATAAATAGTATCGGGTTGGGGACTAAGCGT<br>AAAATCAGCAAAATAGAGTCATAATCCTATGTCTTAGAAATTT<br>TTGAGTTTCTTAGGGCTCAGAGCAGATTAGGTGAAGGCTTATA<br>AATAGTATCGGGTTGGGGACTAAGCGTAAAATCAGCAAAATA<br>GAGTCATAATCCTATGTCTTAGAAATTTTTGAGTTTCTTAGTCA<br>GCAGAGCAGATTAGGTGAAGGCTTATAAATAGTATCGGGTTG<br>GGGACTAAGCGTAAAATCAGCAAAATAGAGTCATAATCCTAT<br>GTCTTAGAAATTTTTGAGTTTCTTAGGTTT |
| HNRNPA<br>2B1-2 | CACCTCTCAAAGTTTTGAAAAGCTATTAGCCAGGATCATGGTG<br>TAATAAGACATAACGTTTTTCCTTTAAAAAAATTTAAGTGCGT<br>GTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATAAAA<br>TAATTCTAAAGGAAATTGTGTAAAACATCTCAAAGTTTTGAAA<br>AGCTATTAGCCAGGATCATGGTGTAATAAGACATAACGTTTTT<br>CCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGCTGT<br>TGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATTGTG<br>TAAGGCTTCTCAAAGTTTTGAAAAGCTATTAGCCAGGATCATG<br>GTGTAATAAGACATAACGTTTTTCCTTTAAAAAAATTTAAGTG<br>CGTGTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATA<br>AAATAATTCTAAAGGAAATTGTGTAATCAGTCTCAAAGTTTTG<br>AAAAGCTATTAGCCAGGATCATGGTGTAATAAGACATAACGTT<br>TTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGC<br>TGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATT<br>GTGTAAGTTT |
| mHNRNP<br>A2B1-2 | CACCTCTCAAAGTTCCAATTAATAGTGCTCTACCTGGTAATAA<br>AATATTGCGTAGAGTTAGTCATGTTAAAAAAATTTAAGTGCGT<br>GTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATAAAA<br>TAATTCTAAAGGAAATTGTGTAAAACATCTCAAAGTTCCAATT<br>AATAGTGCTCTACCTGGTAATAAAATATTGCGTAGAGTTAGTC<br>ATGTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGCTGT<br>TGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATTGTG<br>TAAGGCTTCTCAAAGTTCCAATTAATAGTGCTCTACCTGGTAA<br>TAAAATATTGCGTAGAGTTAGTCATGTTAAAAAAATTTAAGTG<br>CGTGTGTAGAGTTAAGAAGCTGTTGTACATTTATGATTTAATA<br>AAATAATTCTAAAGGAAATTGTGTAATCAGTCTCAAAGTTCCA<br>ATTAATAGTGCTCTACCTGGTAATAAAATATTGCGTAGAGTTA<br>GTCATGTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGC<br>TGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATT<br>GTGTAAGTTT |

| | |
|---|---|
| HNRNPA2B1-3 | CACCGTTTTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTT<br>AAGAAGCTGTTGTACATTTATGATTTAATAAAATAATTCTAAA<br>GGAAATTGTGTAATTATAGACTTTTTATTTTAAATAAGTTAAG<br>GAGTGGGTAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTAT<br>ATTTGTATAAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCT<br>GCAATTCATCAGGATTAAATTATGTAGATAACTTAAGGGAAAC<br>AGTTTTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAG<br>AAGCTGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGA<br>AATTGTGTAATTATAGACTTTTTATTTTAAATAAGTTAAGGAGT<br>GGGTAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTT<br>GTATAAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAA<br>TTCATCAGGATTAAATTATGTAGATAACTTAAGGGAGGCTGTT<br>TTTCCTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGC<br>TGTTGTACATTTATGATTTAATAAAATAATTCTAAAGGAAATT<br>GTGTAATTATAGACTTTTTATTTTAAATAAGTTAAGGAGTGGG<br>TAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTTGTAT<br>AAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAATTCA<br>TCAGGATTAAATTATGTAGATAACTTAAGGGATCAGGTTTTTC<br>CTTTAAAAAAATTTAAGTGCGTGTGTAGAGTTAAGAAGCTGTT<br>GTACATTTATGATTTAATAAAATAATTCTAAAGGAAATTGTGT<br>AATTATAGACTTTTTATTTTAAATAAGTTAAGGAGTGGGTAGT<br>ATAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTTGTATAAG<br>ATAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAATTCATCA<br>GGATTAAATTATGTAGATAACTTAAGGGAGTTT |
| mHNRNPA2B1-3 | CACCGTTTTTCCTTATGTATTTGAAGGGTATGAGAATACATTA<br>GTGGGTTTATTTTTTACTAAAAAGACAATAAAATAATTCTAAA<br>GGAAATTGTGTAATTATAGACTTTTTATTTTAAATAAGTTAAG<br>GAGTGGGTAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTAT<br>ATTTGTATAAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCT<br>GCAATTCATCAGGATTAAATTATGTAGATAACTTAAGGGAAAC<br>AGTTTTTCCTTATGTATTTGAAGGGTATGAGAATACATTAGTG<br>GGTTTATTTTTTACTAAAAAGACAATAAAATAATTCTAAAGGA<br>AATTGTGTAATTATAGACTTTTTATTTTAAATAAGTTAAGGAGT<br>GGGTAGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTT<br>GTATAAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAA<br>TTCATCAGGATTAAATTATGTAGATAACTTAAGGGAGGCTGTT<br>TTTCCTTATGTATTTGAAGGGTATGAGAATACATTAGTGGGTTT<br>ATTTTTTACTAAAAAGACAATAAAATAATTCTAAAGGAAATTG<br>TGTAATTATAGACTTTTTATTTTAAATAAGTTAAGGAGTGGGT<br>AGTATAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTTGTAT<br>AAGATAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAATTCA<br>TCAGGATTAAATTATGTAGATAACTTAAGGGATCAGGTTTTTC<br>CTTATGTATTTGAAGGGTATGAGAATACATTAGTGGGTTTATTT<br>TTTACTAAAAAGACAATAAAATAATTCTAAAGGAAATTGTGTA |

| | |
|---|---|
| | ATTATAGACTTTTTATTTTAAATAAGTTAAGGAGTGGGTAGTA<br>TAATTAAGGTCCGTTGCAAAGCTGTTGTTATATTTGTATAAGA<br>TAAATGCTGGTCAGATGTAAGTGTGTTGTCTGCAATTCATCAG<br>GATTAAATTATGTAGATAACTTAAGGGAGTTT |
| KIAA0907 | CACCGTTTGAGATATTGAACTGTCATTTTTGCACATTTGAATAC<br>TTTGCAGGCTGGCTTTGTATAAACTTATCCTCTGGTTTCCTATA<br>TGTTGTAAATATTTAGACCATAATTTCATTATAAATAAATCTAT<br>AAATATTCTGCTTGTGGTTAACAGTTTGAGATATTGAACTGTC<br>ATTTTTGCACATTTGAATACTTTGCAGGCTGGCTTTGTATAAAC<br>TTATCCTCTGGTTTCCTATATGTTGTAAATATTTAGACCATAAT<br>TTCATTATAAATAAATCTATAAATATTCTGCTTGTGGTTGGCTG<br>TTTGAGATATTGAACTGTCATTTTTGCACATTTGAATACTTTGC<br>AGGCTGGCTTTGTATAAACTTATCCTCTGGTTTCCTATATGTTG<br>TAAATATTTAGACCATAATTTCATTATAAATAAATCTATAAAT<br>ATTCTGCTTGTGGTTTCAGGTTTGAGATATTGAACTGTCATTTT<br>TGCACATTTGAATACTTTGCAGGCTGGCTTTGTATAAACTTATC<br>CTCTGGTTTCCTATATGTTGTAAATATTTAGACCATAATTTCAT<br>TATAAATAAATCTATAAATATTCTGCTTGTGGTTGTTT |
| mKIAA0907 | CACCGTTTGAGATATAGTGTTCGGATTATTGTCCTATCTTCGTG<br>ATATTGATTTCGATCGCAATAAACTTATCCTCTGGTTTCCTATA<br>TGTTGTAAATATTTAGACCATAATTTCATTATAAATAAATCTAT<br>AAATATTCTGCTTGTGGTTAACAGTTTGAGATATAGTGTTCGG<br>ATTATTGTCCTATCTTCGTGATATTGATTTCGATCGCAATAAAC<br>TTATCCTCTGGTTTCCTATATGTTGTAAATATTTAGACCATAAT<br>TTCATTATAAATAAATCTATAAATATTCTGCTTGTGGTTGGCTG<br>TTTGAGATATAGTGTTCGGATTATTGTCCTATCTTCGTGATATT<br>GATTTCGATCGCAATAAACTTATCCTCTGGTTTCCTATATGTTG<br>TAAATATTTAGACCATAATTTCATTATAAATAAATCTATAAAT<br>ATTCTGCTTGTGGTTTCAGGTTTGAGATATAGTGTTCGGATTAT<br>TGTCCTATCTTCGTGATATTGATTTCGATCGCAATAAACTTATC<br>CTCTGGTTTCCTATATGTTGTAAATATTTAGACCATAATTTCAT<br>TATAAATAAATCTATAAATATTCTGCTTGTGGTTGTTT |
| KLHL15 | CACCAATGTTACTGGTTTTATCTACTTGTTTATTTTGTACAAAA<br>TACCCAGCGACACTAGGGATGTAAGCCCTCAGTTTTAACAAAT<br>GTTACTGGTTTTATCTACTTGTTTATTTTGTACAAAATACCCAG<br>CGACACTAGGGATGTAAGCCCTCAGTTTTGGCTAATGTTACTG<br>GTTTTATCTACTTGTTTATTTTGTACAAAATACCCAGCGACACT<br>AGGGATGTAAGCCCTCAGTTTTTCAGAATGTTACTGGTTTTATC<br>TACTTGTTTATTTTGTACAAAATACCCAGCGACACTAGGGATG<br>TAAGCCCTCAGTTTTGTTT |
| mKLHL15 | CACCAATGTTACTGGTTTTATTAAGATTTAGTACGTCCCAGTA<br>AGTCCGAATGTCCTTATCGACATGTAACCTCAGTTTTAACAAA<br>TGTTACTGGTTTTATTAAGATTTAGTACGTCCCAGTAAGTCCGA<br>ATGTCCTTATCGACATGTAACCTCAGTTTTGGCTAATGTTACTG |

| | |
|---|---|
| | GTTTTATTAAGATTTAGTACGTCCCAGTAAGTCCGAATGTCCTT<br>ATCGACATGTAACCTCAGTTTTTCAGAATGTTACTGGTTTTATT<br>AAGATTTAGTACGTCCCAGTAAGTCCGAATGTCCTTATCGACA<br>TGTAACCTCAGTTTTGTTT |
| PMEPA1 | CACCGTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCA<br>CTTAACTTTAAAGACATATTTGCACAAAACCTTTGTTTAAAGA<br>TCTGCAATATTATATAACAGTGCGTGAATGCTTATTTTCTTTTG<br>TTTATGATAATTTCACTTAACTTTAAAGACATATTTGCACAAA<br>ACCTTTGTTTAAAGATCTGCAATATTATATGGCTGTGCGTGAA<br>TGCTTATTTTCTTTTGTTTATGATAATTTCACTTAACTTTAAAGA<br>CATATTTGCACAAAACCTTTGTTTAAAGATCTGCAATATTATAT<br>TCAGGTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCA<br>CTTAACTTTAAAGACATATTTGCACAAAACCTTTGTTTAAAGA<br>TCTGCAATATTATATGTTT |
| mPMEPA1 | CACCGTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCA<br>CTTAATGCTTATAGCTATACCTTATGAACCCGTTAACATTTGAA<br>ATAAAATATTATATAACAGTGCGTGAATGCTTATTTTCTTTTGT<br>TTATGATAATTTCACTTAATGCTTATAGCTATACCTTATGAACC<br>CGTTAACATTTGAAATAAAATATTATATGGCTGTGCGTGAATG<br>CTTATTTTCTTTTGTTTATGATAATTTCACTTAATGCTTATAGCT<br>ATACCTTATGAACCCGTTAACATTTGAAATAAAATATTATATT<br>CAGGTGCGTGAATGCTTATTTTCTTTTGTTTATGATAATTTCAC<br>TTAATGCTTATAGCTATACCTTATGAACCCGTTAACATTTGAA<br>ATAAAATATTATATGTTT |
| PNRC1 | CACCATACAAACAGCTTGTATTATATTTTATATTTTGTAAATAC<br>TGTATACCATGTATTATGTGTATATTGTTCATACTTGAGAGGTA<br>TATTATAGTTTTGTTATGAAAGTATGTATTTTGCCCTGCCCACA<br>TTGCAGGTGTTTTGTATATATACAATGGATAAATTTTAAGTGTG<br>TGCTAAGGAACAATACAAACAGCTTGTATTATATTTTATATTTT<br>GTAAATACTGTATACCATGTATTATGTGTATATTGTTCATACTT<br>GAGAGGTATATTATAGTTTTGTTATGAAAGTATGTATTTTGCCC<br>TGCCCACATTGCAGGTGTTTTGTATATATACAATGGATAAATT<br>TTAAGTGTGTGCTAAGGGGCTATACAAACAGCTTGTATTATAT<br>TTTATATTTTGTAAATACTGTATACCATGTATTATGTGTATATT<br>GTTCATACTTGAGAGGTATATTATAGTTTTGTTATGAAAGTATG<br>TATTTTGCCCTGCCCACATTGCAGGTGTTTTGTATATATACAAT<br>GGATAAATTTTAAGTGTGTGCTAAGGTCAGATACAAACAGCTT<br>GTATTATATTTTATATTTTGTAAATACTGTATACCATGTATTAT<br>GTGTATATTGTTCATACTTGAGAGGTATATTATAGTTTTGTTAT<br>GAAAGTATGTATTTTGCCCTGCCCACATTGCAGGTGTTTTGTAT<br>ATATACAATGGATAAATTTTAAGTGTGTGCTAAGGGTTT |
| mPNRC1 | CACCATACAAACAGCTTGTATTATATTTTATATTTTGTAAATAC<br>TGTATACCATGTATTCGATTGTTATTGTTAGGTATAGTGTACAA<br>TATTATAGTTTTGTTATGAAAGTATGTATTTTGCCCTGCCCACA |

| | |
|---|---|
| | TTGCAGGTGTTTTGTATATATACAATGGATAAATTTTAAGTGTG<br>TGCTAAGGAACAATACAAACAGCTTGTATTATATTTTATATTTT<br>GTAAATACTGTATACCATGTATTCGATTGTTATTGTTAGGTATA<br>GTGTACAATATTATAGTTTTGTTATGAAAGTATGTATTTTGCCC<br>TGCCCACATTGCAGGTGTTTTGTATATATACAATGGATAAATT<br>TTAAGTGTGTGCTAAGGGGCTATACAAACAGCTTGTATTATAT<br>TTTATATTTTGTAAATACTGTATACCATGTATTCGATTGTTATT<br>GTTAGGTATAGTGTACAATATTATAGTTTTGTTATGAAAGTAT<br>GTATTTTGCCCTGCCCACATTGCAGGTGTTTTGTATATATACAA<br>TGGATAAATTTTAAGTGTGTGCTAAGGTCAGATACAAACAGCT<br>TGTATTATATTTTATATTTTGTAAATACTGTATACCATGTATTC<br>GATTGTTATTGTTAGGTATAGTGTACAATATTATAGTTTTGTTA<br>TGAAAGTATGTATTTTGCCCTGCCCACATTGCAGGTGTTTTGTA<br>TATATACAATGGATAAATTTTAAGTGTGTGCTAAGGGTTT |
| RGMA | CACCCTGCGTCCACGTGTCTGCGACCTGTGTGGAGTGTCACCG<br>CGTGTACATACTGTAAATTATTTATTAATGGCTAAATGCAAAA<br>CACTGCGTCCACGTGTCTGCGACCTGTGTGGAGTGTCACCGCG<br>TGTACATACTGTAAATTATTTATTAATGGCTAAATGCAAGGCT<br>CTGCGTCCACGTGTCTGCGACCTGTGTGGAGTGTCACCGCGTG<br>TACATACTGTAAATTATTTATTAATGGCTAAATGCAATCAGCT<br>GCGTCCACGTGTCTGCGACCTGTGTGGAGTGTCACCGCGTGTA<br>CATACTGTAAATTATTTATTAATGGCTAAATGCAAGTTT |
| mRGMA | CACCCTGCGTCCACAGGGGGGTGCGTTGTCACGTTTATGTTCT<br>ACGGCAGCCCTACTAAATTATTTATTAATGGCTAAATGCAAAA<br>CACTGCGTCCACAGGGGGGTGCGTTGTCACGTTTATGTTCTAC<br>GGCAGCCCTACTAAATTATTTATTAATGGCTAAATGCAAGGCT<br>CTGCGTCCACAGGGGGGTGCGTTGTCACGTTTATGTTCTACGG<br>CAGCCCTACTAAATTATTTATTAATGGCTAAATGCAATCAGCT<br>GCGTCCACAGGGGGGTGCGTTGTCACGTTTATGTTCTACGGCA<br>GCCCTACTAAATTATTTATTAATGGCTAAATGCAAGTTT |
| RPA2 | CACCACTTTTTGACACACTTGCCATGACGTGTGTTTCTGTGAAC<br>ATGAAGTTCTGCGGTAGTGCCTCCAGGGGCAGAGGAAAAGAA<br>GAAGTGTTACTGCATTTTGTACAAAATAAATACAGTCATATGT<br>TTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGTT<br>TAACAACTTTTTGACACACTTGCCATGACGTGTGTTTCTGTGAA<br>CATGAAGTTCTGCGGTAGTGCCTCCAGGGGCAGAGGAAAAGA<br>AGAAGTGTTACTGCATTTTGTACAAAATAAATACAGTCATATG<br>TTTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGT<br>TTGGCTACTTTTTGACACACTTGCCATGACGTGTGTTTCTGTGA<br>ACATGAAGTTCTGCGGTAGTGCCTCCAGGGGCAGAGGAAAAG<br>AAGAAGTGTTACTGCATTTTGTACAAAATAAATACAGTCATAT<br>GTTTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTC<br>GTTTTCAGACTTTTTGACACACTTGCCATGACGTGTGTTTCTGT<br>GAACATGAAGTTCTGCGGTAGTGCCTCCAGGGGCAGAGGAAA |

| | |
|---|---|
| | AGAAGAAGTGTTACTGCATTTTGTACAAAATAAATACAGTCAT<br>ATGTTTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTC<br>TCGTTTGTTT |
| mRPA2 | CACCACTTTTTGACGCCTATAGGTGCTGTTCGCCGTTTAGGTAG<br>ATCCATTCTAGGAGGATAATATGAAGAAAAATCTACGTCGCTG<br>AATCATCAATCGGATGGTGGGATGCAAAATACAGTCATATGTT<br>TAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGTTT<br>AACAACTTTTTGACGCCTATAGGTGCTGTTCGCCGTTTAGGTA<br>GATCCATTCTAGGAGGATAATATGAAGAAAAATCTACGTCGCT<br>GAATCATCAATCGGATGGTGGGATGCAAAATACAGTCATATGT<br>TTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGTT<br>TGGCTACTTTTTGACGCCTATAGGTGCTGTTCGCCGTTTAGGTA<br>GATCCATTCTAGGAGGATAATATGAAGAAAAATCTACGTCGCT<br>GAATCATCAATCGGATGGTGGGATGCAAAATACAGTCATATGT<br>TTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGTT<br>TTCAGACTTTTTGACGCCTATAGGTGCTGTTCGCCGTTTAGGTA<br>GATCCATTCTAGGAGGATAATATGAAGAAAAATCTACGTCGCT<br>GAATCATCAATCGGATGGTGGGATGCAAAATACAGTCATATGT<br>TTAATAAAACAGTTCTATTGTAGTAACTTGTAAAAATTCTCGTT<br>TGTTT |
| RRAGD | CACCCTTCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAA<br>ACTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCATTA<br>CACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGT<br>GTGAGAACTGCCAGGAATTGTACATATTTTGCACTTAACACTT<br>CTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAAACTTGTA<br>TATTAACAGTGATGTGAAATCCATTGTCATTTCATTACACAAA<br>TGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGA<br>ACTGCCAGGAATTGTACATATTTTGCACTTGGCTCTTCTCTTTT<br>ATAAATAAAGTAAGCACTTTGAAGCAAAAACTTGTATATTAAC<br>AGTGATGTGAAATCCATTGTCATTTCATTACACAAATGTAAAC<br>TTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCA<br>GGAATTGTACATATTTTGCACTTTCAGCTTCTCTTTTATAAATA<br>AAGTAAGCACTTTGAAGCAAAAACTTGTATATTAACAGTGATG<br>TGAAATCCATTGTCATTTCATTACACAAATGTAAACTTTTATGG<br>TCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCAGGAATTGT<br>ACATATTTTGCACTTGTTT |
| mRRAGD | CACCCTTCTCTTTTAATTAGTTGAACCAAGACATAGAATAAAA<br>ACTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCATTA<br>CACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGT<br>GTGAGAACTGCCAGGAATTGTACATATTTTGCACTTAACACTT<br>CTCTTTTAATTAGTTGAACCAAGACATAGAATAAAAACTTGTA<br>TATTAACAGTGATGTGAAATCCATTGTCATTTCATTACACAAA<br>TGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGA<br>ACTGCCAGGAATTGTACATATTTTGCACTTGGCTCTTCTCTTTT |

| | |
|---|---|
| | AATTAGTTGAACCAAGACATAGAATAAAAACTTGTATATTAAC<br>AGTGATGTGAAATCCATTGTCATTTCATTACACAAATGTAAAC<br>TTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCA<br>GGAATTGTACATATTTTGCACTTTCAGCTTCTCTTTTAATTAGT<br>TGAACCAAGACATAGAATAAAAACTTGTATATTAACAGTGAT<br>GTGAAATCCATTGTCATTTCATTACACAAATGTAAACTTTTATG<br>GTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCAGGAATTG<br>TACATATTTTGCACTTGTTT |
| mmiR_RRAGD | CACCCTTCTCTTTTATAAATAAAGTAAGGTGTTTGAAGCAAAA<br>ACTTGTATATTAACAGTGATGTGAAATCCATTGTCATTTCATTA<br>CACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGT<br>GTGAGAACTGCCAGGAATTGTACATATTTTGCACTTAACACTT<br>CTCTTTTATAAATAAAGTAAGGTGTTTGAAGCAAAAACTTGTA<br>TATTAACAGTGATGTGAAATCCATTGTCATTTCATTACACAAA<br>TGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGA<br>ACTGCCAGGAATTGTACATATTTTGCACTTGGCTCTTCTCTTTT<br>ATAAATAAAGTAAGGTGTTTGAAGCAAAAACTTGTATATTAAC<br>AGTGATGTGAAATCCATTGTCATTTCATTACACAAATGTAAAC<br>TTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCA<br>GGAATTGTACATATTTTGCACTTTCAGCTTCTCTTTTATAAATA<br>AAGTAAGGTGTTTGAAGCAAAAACTTGTATATTAACAGTGATG<br>TGAAATCCATTGTCATTTCATTACACAAATGTAAACTTTTATGG<br>TCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCAGGAATTGT<br>ACATATTTTGCACTTGTTT |
| mP_RRAGD | CACCCTTCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAA<br>ACTATAATATTAACAGTGATGTGAAATCCATTGTCATTTCATT<br>ACACAAATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCG<br>TGTGAGAACTGCCAGGAATATAACATATTTTGCACTTAACACT<br>TCTCTTTTATAAATAAAGTAAGCACTTTGAAGCAAAAACTATA<br>ATATTAACAGTGATGTGAAATCCATTGTCATTTCATTACACAA<br>ATGTAAACTTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAG<br>AACTGCCAGGAATATAACATATTTTGCACTTGGCTCTTCTCTTT<br>TATAAATAAAGTAAGCACTTTGAAGCAAAAACTATAATATTAA<br>CAGTGATGTGAAATCCATTGTCATTTCATTACACAAATGTAAA<br>CTTTTATGGTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCC<br>AGGAATATAACATATTTTGCACTTTCAGCTTCTCTTTTATAAAT<br>AAAGTAAGCACTTTGAAGCAAAAACTATAATATTAACAGTGA<br>TGTGAAATCCATTGTCATTTCATTACACAAATGTAAACTTTTAT<br>GGTCTGTAGTCAAAAAAATCCCGTGTGAGAACTGCCAGGAAT<br>ATAACATATTTTGCACTTGTTT |
| SNAPC1 | CACCATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTA<br>TTTTGTATATTTATTCTATTTTGTATATTGTTAAGTGCAATAAA<br>GTTTTTGCCTTGCTAACAATAACTATTTTGTATCTACAGTCGGA<br>TAATGGATTTTTTATTTTGTATATTTATTCTATTTTGTATATTGT |

| | |
|---|---|
| | TAAGTGCAATAAAGTTTTTGCCTTGCTGGCTATAACTATTTTGT<br>ATCTACAGTCGGATAATGGATTTTTTATTTTGTATATTTATTCT<br>ATTTTGTATATTGTTAAGTGCAATAAAGTTTTTGCCTTGCTTCA<br>GATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTATTT<br>TGTATATTTATTCTATTTTGTATATTGTTAAGTGCAATAAAGTT<br>TTTGCCTTGCTGTTT |
| mSNAPC1 | CACCATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTA<br>TTTTGTATATTTATTCTATTTTGTACAGAAATTTTATGATGTAA<br>TTGTTTGCCTTGCTAACAATAACTATTTTGTATCTACAGTCGGA<br>TAATGGATTTTTTATTTTGTATATTTATTCTATTTTGTACAGAA<br>ATTTTATGATGTAATTGTTTGCCTTGCTGGCTATAACTATTTTG<br>TATCTACAGTCGGATAATGGATTTTTTATTTTGTATATTTATTC<br>TATTTTGTACAGAAATTTTATGATGTAATTGTTTGCCTTGCTTC<br>AGATAACTATTTTGTATCTACAGTCGGATAATGGATTTTTTATT<br>TTGTATATTTATTCTATTTTGTACAGAAATTTTATGATGTAATT<br>GTTTGCCTTGCTGTTT |
| TOB1 | CACCAGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGT<br>AAATATTAATTTTGTACCTATATTGTGCAATACTTGAAAAAAA<br>CGGTATAAAAGAACAAGATTTTTGCTATATATTATGGAAGAAA<br>AATGTAATCGTAAATATTAATTTTGTACCTATATTGTGCAATAC<br>TTGAAAAAAACGGTATAAAAGGGCTAGATTTTTGCTATATATT<br>ATGGAAGAAAAATGTAATCGTAAATATTAATTTTGTACCTATA<br>TTGTGCAATACTTGAAAAAAACGGTATAAAAGTCAGAGATTTT<br>TGCTATATATTATGGAAGAAAAATGTAATCGTAAATATTAATT<br>TTGTACCTATATTGTGCAATACTTGAAAAAAACGGTATAAAAG<br>GTTT |
| mTOB1 | CACCAGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGT<br>AAATAATATTCTTACAGGAAATCGAACTAATGTATTAATTTCA<br>TGGTATAAAAGAACAAGATTTTTGCTATATATTATGGAAGAAA<br>AATGTAATCGTAAATAATATTCTTACAGGAAATCGAACTAATG<br>TATTAATTTCATGGTATAAAAGGGCTAGATTTTTGCTATATATT<br>ATGGAAGAAAAATGTAATCGTAAATAATATTCTTACAGGAAA<br>TCGAACTAATGTATTAATTTCATGGTATAAAAGTCAGAGATTT<br>TTGCTATATATTATGGAAGAAAAATGTAATCGTAAATAATATT<br>CTTACAGGAAATCGAACTAATGTATTAATTTCATGGTATAAAA<br>GGTTT |
| mmiR_TOB1 | CACCAGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGT<br>AAATATTAATTTTGTACCTATATTGACGAATACTTGAAAAAAA<br>CGGTATAAAAGAACAAGATTTTTGCTATATATTATGGAAGAAA<br>AATGTAATCGTAAATATTAATTTTGTACCTATATTGACGAATA<br>CTTGAAAAAAACGGTATAAAAGGGCTAGATTTTTGCTATATAT<br>TATGGAAGAAAAATGTAATCGTAAATATTAATTTTGTACCTAT<br>ATTGACGAATACTTGAAAAAAACGGTATAAAAGTCAGAGATT<br>TTTGCTATATATTATGGAAGAAAAATGTAATCGTAAATATTAA |

| | |
|---|---|
| | TTTTGTACCTATATTGACGAATACTTGAAAAAAACGGTATAAA<br>AGGTTT |
| mP_TOB1 | CACCAGATTTTTGCTATATATTATGGAAGAAAAATGTAATCGT<br>AAATATTAATTTATAACCTATATTGTGCAATACTTGAAAAAAA<br>CGGTATAAAAGAACAAGATTTTTGCTATATATTATGGAAGAAA<br>AATGTAATCGTAAATATTAATTTATAACCTATATTGTGCAATA<br>CTTGAAAAAAACGGTATAAAAGGGCTAGATTTTTGCTATATAT<br>TATGGAAGAAAAATGTAATCGTAAATATTAATTTATAACCTAT<br>ATTGTGCAATACTTGAAAAAAACGGTATAAAAGTCAGAGATTT<br>TTGCTATATATTATGGAAGAAAAATGTAATCGTAAATATTAAT<br>TTATAACCTATATTGTGCAATACTTGAAAAAAACGGTATAAAA<br>GGTTT |
| VLDLR | CACCCTTGACCGTTTTTATATTACTTTTGTAAATATTCTTGTCC<br>ACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAAC<br>CCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGC<br>ACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCTA<br>TAGTGTATACCACCTGTACATACATTGTATAGGCCATCTGTAA<br>ATATCCCAGAGAACAATCACTATTCTTAAGCACTTTGAAAATA<br>TTTCTATGTAAATTATTGTAAACTTTTTCAATGGTTGGGACAAT<br>GGCAATAGGACAAAACGGGTTACTAAGATGAACACTTGACCG<br>TTTTTATATTACTTTTGTAAATATTCTTGTCCACATTCTACTTCA<br>GCTTTGGATGTGGTTACCGAGTATCTGTAACCCTTGAATTTCTA<br>GACAGTATTGCCACCTCTGGCCAAATATGCACTTTCCCTAGAA<br>AGCCATATTCCAGCAGTGAAACTTGTGCTATAGTGTATACCAC<br>CTGTACATACATTGTATAGGCCATCTGTAAATATCCCAGAGAA<br>CAATCACTATTCTTAAGCACTTTGAAAATATTTCTATGTAAATT<br>ATTGTAAACTTTTTCAATGGTTGGGACAATGGCAATAGGACAA<br>AACGGGTTACTAAGATGGGCTCTTGACCGTTTTTATATTACTTT<br>TGTAAATATTCTTGTCCACATTCTACTTCAGCTTTGGATGTGGT<br>TACCGAGTATCTGTAACCCTTGAATTTCTAGACAGTATTGCCA<br>CCTCTGGCCAAATATGCACTTTCCCTAGAAAGCCATATTCCAG<br>CAGTGAAACTTGTGCTATAGTGTATACCACCTGTACATACATT<br>GTATAGGCCATCTGTAAATATCCCAGAGAACAATCACTATTCT<br>TAAGCACTTTGAAAATATTTCTATGTAAATTATTGTAAACTTTT<br>TCAATGGTTGGGACAATGGCAATAGGACAAAACGGGTTACTA<br>AGATGTCAGCTTGACCGTTTTTATATTACTTTTGTAAATATTCT<br>TGTCCACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCT<br>GTAACCCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAA<br>TATGCACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTG<br>TGCTATAGTGTATACCACCTGTACATACATTGTATAGGCCATC<br>TGTAAATATCCCAGAGAACAATCACTATTCTTAAGCACTTTGA<br>AAATATTTCTATGTAAATTATTGTAAACTTTTTCAATGGTTGGG<br>ACAATGGCAATAGGACAAAACGGGTTACTAAGATGGTTT |

| | |
|---|---|
| mVLDLR | CACCCTTGACCGTTTTTATATTACTTTTGTAAATATTCTTGTCC<br>ACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAAC<br>CCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGC<br>ACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCTA<br>TAGTGTATACCACCTGTACATACATTGTATAGGCCATCTGTAA<br>ATATCCCTTAAAGTGTCGAACCTCAAATATTAATACAAACGTA<br>TTTCTATGTAAATTATTGTAAACTTTTTCAATGGTTGGGACAAT<br>GGCAATAGGACAAAACGGGTTACTAAGATGAACACTTGACCG<br>TTTTTATATTACTTTTGTAAATATTCTTGTCCACATTCTACTTCA<br>GCTTTGGATGTGGTTACCGAGTATCTGTAACCCTTGAATTTCTA<br>GACAGTATTGCCACCTCTGGCCAAATATGCACTTTCCCTAGAA<br>AGCCATATTCCAGCAGTGAAACTTGTGCTATAGTGTATACCAC<br>CTGTACATACATTGTATAGGCCATCTGTAAATATCCCTTAAAG<br>TGTCGAACCTCAAATATTAATACAAACGTATTTCTATGTAAAT<br>TATTGTAAACTTTTTCAATGGTTGGGACAATGGCAATAGGACA<br>AAACGGGTTACTAAGATGGGCTCTTGACCGTTTTTATATTACTT<br>TTGTAAATATTCTTGTCCACATTCTACTTCAGCTTTGGATGTGG<br>TTACCGAGTATCTGTAACCCTTGAATTTCTAGACAGTATTGCC<br>ACCTCTGGCCAAATATGCACTTTCCCTAGAAAGCCATATTCCA<br>GCAGTGAAACTTGTGCTATAGTGTATACCACCTGTACATACAT<br>TGTATAGGCCATCTGTAAATATCCCTTAAAGTGTCGAACCTCA<br>AATATTAATACAAACGTATTTCTATGTAAATTATTGTAAACTTT<br>TTCAATGGTTGGGACAATGGCAATAGGACAAAACGGGTTACT<br>AAGATGTCAGCTTGACCGTTTTTATATTACTTTTGTAAATATTC<br>TTGTCCACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATC<br>TGTAACCCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAA<br>ATATGCACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTT<br>GTGCTATAGTGTATACCACCTGTACATACATTGTATAGGCCAT<br>CTGTAAATATCCCTTAAAGTGTCGAACCTCAAATATTAATACA<br>AACGTATTTCTATGTAAATTATTGTAAACTTTTTCAATGGTTGG<br>GACAATGGCAATAGGACAAAACGGGTTACTAAGATGGTTT |
| mP_VLDLR | CACCCTTGACCGTTTTTATATTACTTTATAAAATATTCTTGTCC<br>ACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCTGTAAC<br>CCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAATATGC<br>ACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTGTGCTA<br>TAGTGTATACCACCATAACATACATTGTATAGGCCATCATAAA<br>ATATCCCAGAGAACAATCACTATTCTTAAGCACTTTGAAAATA<br>TTTCTAATAAAATTATTGTAAACTTTTTCAATGGTTGGGACAAT<br>GGCAATAGGACAAAACGGGTTACTAAGATGAACACTTGACCG<br>TTTTTATATTACTTTATAAAATATTCTTGTCCACATTCTACTTCA<br>GCTTTGGATGTGGTTACCGAGTATCTGTAACCCTTGAATTTCTA<br>GACAGTATTGCCACCTCTGGCCAAATATGCACTTTCCCTAGAA<br>AGCCATATTCCAGCAGTGAAACTTGTGCTATAGTGTATACCAC<br>CATAACATACATTGTATAGGCCATCATAAAATATCCCAGAGAA |

| | |
|---|---|
| | CAATCACTATTCTTAAGCACTTTGAAAATATTTCTAATAAAATT<br>ATTGTAAACTTTTTCAATGGTTGGGACAATGGCAATAGGACAA<br>AACGGGTTACTAAGATGGGCTCTTGACCGTTTTTATATTACTTT<br>ATAAAATATTCTTGTCCACATTCTACTTCAGCTTTGGATGTGGT<br>TACCGAGTATCTGTAACCCTTGAATTTCTAGACAGTATTGCCA<br>CCTCTGGCCAAATATGCACTTTCCCTAGAAAGCCATATTCCAG<br>CAGTGAAACTTGTGCTATAGTGTATACCACCATAACATACATT<br>GTATAGGCCATCATAAAATATCCCAGAGAACAATCACTATTCT<br>TAAGCACTTTGAAAATATTTCTAATAAAATTATTGTAAACTTTT<br>TCAATGGTTGGGACAATGGCAATAGGACAAAACGGGTTACTA<br>AGATGTCAGCTTGACCGTTTTTATATTACTTTATAAAATATTCT<br>TGTCCACATTCTACTTCAGCTTTGGATGTGGTTACCGAGTATCT<br>GTAACCCTTGAATTTCTAGACAGTATTGCCACCTCTGGCCAAA<br>TATGCACTTTCCCTAGAAAGCCATATTCCAGCAGTGAAACTTG<br>TGCTATAGTGTATACCACCATAACATACATTGTATAGGCCATC<br>ATAAAATATCCCAGAGAACAATCACTATTCTTAAGCACTTTGA<br>AAATATTTCTAATAAAATTATTGTAAACTTTTTCAATGGTTGGG<br>ACAATGGCAATAGGACAAAACGGGTTACTAAGATGGTTT |
| ZC3H12C | CACCATTTCTTGATACTGCACTATAGAGAAATGGTGATGGAGG<br>AGTTGTAAATGGTAACTTAAAATTTTTGTAAGATATTGTATATT<br>TTCCATTTTCCTGAAGGTAGTTTTCTTGGGGGGGCCTGTTATAT<br>TATTAAGAACAATTTCTTGATACTGCACTATAGAGAAATGGTG<br>ATGGAGGAGTTGTAAATGGTAACTTAAAATTTTTGTAAGATAT<br>TGTATATTTTCCATTTTCCTGAAGGTAGTTTTCTTGGGGGGGCC<br>TGTTATATTATTAAGGGCTATTTCTTGATACTGCACTATAGAGA<br>AATGGTGATGGAGGAGTTGTAAATGGTAACTTAAAATTTTTGT<br>AAGATATTGTATATTTTCCATTTTCCTGAAGGTAGTTTTCTTGG<br>GGGGGCCTGTTATATTATTAAGTCAGATTTCTTGATACTGCACT<br>ATAGAGAAATGGTGATGGAGGAGTTGTAAATGGTAACTTAAA<br>ATTTTTGTAAGATATTGTATATTTTCCATTTTCCTGAAGGTAGT<br>TTTCTTGGGGGGGCCTGTTATATTATTAAGGTTT |
| mZC3H12C | CACCATTTCTTGATTTGGGCTAGGGTAATTAATGAATAGAACC<br>ATAATTTGATAGAGCGGGAAAATTTTTGTAAGATATTGTATAT<br>TTTCCATTTTCCTGAAGGTAGTTTTCTTGGGGGGGCCTGTTATA<br>TTATTAAGAACAATTTCTTGATTTGGGCTAGGGTAATTAATGA<br>ATAGAACCATAATTTGATAGAGCGGGAAAATTTTTGTAAGATA<br>TTGTATATTTTCCATTTTCCTGAAGGTAGTTTTCTTGGGGGGGC<br>CTGTTATATTATTAAGGGCTATTTCTTGATTTGGGCTAGGGTAA<br>TTAATGAATAGAACCATAATTTGATAGAGCGGGAAAATTTTTG<br>TAAGATATTGTATATTTTCCATTTTCCTGAAGGTAGTTTTCTTG<br>GGGGGGCCTGTTATATTATTAAGTCAGATTTCTTGATTTGGGCT<br>AGGGTAATTAATGAATAGAACCATAATTTGATAGAGCGGGAA<br>AATTTTTGTAAGATATTGTATATTTTCCATTTTCCTGAAGGTAG<br>TTTTCTTGGGGGGGCCTGTTATATTATTAAGGTTT |

| | |
|---|---|
| ZIC2 | CACCTTATGAGGCAACCTGATTGTAAACTTCATGTAACTATAG<br>ACTGGAAAAAATGAGCCGTGCCAAAGTCTCCCTTCTGTTTCTT<br>CAGCACATTGACCCATAGCACACACATACACACCACCACCAA<br>CAACGCTTGTGAATGTATTTTTCTGTTAGCTGGGTTTACATGTG<br>ATGTTTTAGTGCTTTTGCAAGTTCAATTTGTTAGTTCCTGTATG<br>AAAGATTGTGGGGGAAAAATAAACGTCGTGCCGTTAGCTTTTT<br>CCGTAATAACACCCTTCCTTCTGTAAATACCCGTTACCATATTT<br>ATCCATTTGTAATTAAATTATGGTATTAACTAACATTATGAGG<br>CAACCTGATTGTAAACTTCATGTAACTATAGACTGGAAAAAAT<br>GAGCCGTGCCAAAGTCTCCCTTCTGTTTCTTCAGCACATTGACC<br>CATAGCACACACATACACACCACCACCAACAACGCTTGTGAAT<br>GTATTTTTCTGTTAGCTGGGTTTACATGTGATGTTTTAGTGCTT<br>TTGCAAGTTCAATTTGTTAGTTCCTGTATGAAAGATTGTGGGG<br>GAAAAATAAACGTCGTGCCGTTAGCTTTTTCCGTAATAACACC<br>CTTCCTTCTGTAAATACCCGTTACCATATTTATCCATTTGTAAT<br>TAAATTATGGTATTAACTGGCTTTATGAGGCAACCTGATTGTA<br>AACTTCATGTAACTATAGACTGGAAAAAATGAGCCGTGCCAA<br>AGTCTCCCTTCTGTTTCTTCAGCACATTGACCCATAGCACACAC<br>ATACACACCACCACCAACAACGCTTGTGAATGTATTTTTCTGT<br>TAGCTGGGTTTACATGTGATGTTTTAGTGCTTTTGCAAGTTCAA<br>TTTGTTAGTTCCTGTATGAAAGATTGTGGGGGAAAAATAAACG<br>TCGTGCCGTTAGCTTTTTCCGTAATAACACCCTTCCTTCTGTAA<br>ATACCCGTTACCATATTTATCCATTTGTAATTAAATTATGGTAT<br>TAACTTCAGTTATGAGGCAACCTGATTGTAAACTTCATGTAAC<br>TATAGACTGGAAAAAATGAGCCGTGCCAAAGTCTCCCTTCTGT<br>TTCTTCAGCACATTGACCCATAGCACACACATACACACCACCA<br>CCAACAACGCTTGTGAATGTATTTTTCTGTTAGCTGGGTTTACA<br>TGTGATGTTTTAGTGCTTTTGCAAGTTCAATTTGTTAGTTCCTG<br>TATGAAAGATTGTGGGGGAAAAATAAACGTCGTGCCGTTAGC<br>TTTTTCCGTAATAACACCCTTCCTTCTGTAAATACCCGTTACCA<br>TATTTATCCATTTGTAATTAAATTATGGTATTAACTGTTT |
| mZIC2 | CACCTTATGAGGCAACCTGATTGTAAACTTCATGTCGATAGTT<br>CTTAACCCAACCACCAGCTATTGCGACTCTGACGATCCGATGA<br>TTACAGGCAATATTATAGCACACACATACACACCACCACCAAC<br>AACGCTTGTGAATGTATTTTTCTGTTAGCTGGGTTTACATGTGA<br>TGTTTTAGTGCTTTTGCAAGTTCAATTTGTTAGTTCCTGTATGA<br>AAGATTGTGGGGGAAAAATAAACGTCGTGCCGTTAGCTTTTTC<br>CGTAATAACACCCTTCCTTCTGTAAATACCCGTTACCATATTTA<br>TCCATTTGTAATTAAATTATGGTATTAACTAACATTATGAGGC<br>AACCTGATTGTAAACTTCATGTCGATAGTTCTTAACCCAACCA<br>CCAGCTATTGCGACTCTGACGATCCGATGATTACAGGCAATAT<br>TATAGCACACACATACACACCACCACCAACAACGCTTGTGAAT<br>GTATTTTTCTGTTAGCTGGGTTTACATGTGATGTTTTAGTGCTT<br>TTGCAAGTTCAATTTGTTAGTTCCTGTATGAAAGATTGTGGGG |

| | |
|---|---|
| | GAAAAATAAACGTCGTGCCGTTAGCTTTTTCCGTAATAACACC<br>CTTCCTTCTGTAAATACCCGTTACCATATTTATCCATTTGTAAT<br>TAAATTATGGTATTAACTGGCTTTATGAGGCAACCTGATTGTA<br>AACTTCATGTCGATAGTTCTTAACCCAACCACCAGCTATTGCG<br>ACTCTGACGATCCGATGATTACAGGCAATATTATAGCACACAC<br>ATACACACCACCACCAACAACGCTTGTGAATGTATTTTTCTGT<br>TAGCTGGGTTTACATGTGATGTTTTAGTGCTTTTGCAAGTTCAA<br>TTTGTTAGTTCCTGTATGAAAGATTGTGGGGGAAAAATAAACG<br>TCGTGCCGTTAGCTTTTTCCGTAATAACACCCTTCCTTCTGTAA<br>ATACCCGTTACCATATTTATCCATTTGTAATTAAATTATGGTAT<br>TAACTTCAGTTATGAGGCAACCTGATTGTAAACTTCATGTCGA<br>TAGTTCTTAACCCAACCACCAGCTATTGCGACTCTGACGATCC<br>GATGATTACAGGCAATATTATAGCACACACATACACACCACCA<br>CCAACAACGCTTGTGAATGTATTTTTCTGTTAGCTGGGTTTACA<br>TGTGATGTTTTAGTGCTTTTGCAAGTTCAATTTGTTAGTTCCTG<br>TATGAAAGATTGTGGGGGAAAAATAAACGTCGTGCCGTTAGC<br>TTTTTCCGTAATAACACCCTTCCTTCTGTAAATACCCGTTACCA<br>TATTTATCCATTTGTAATTAAATTATGGTATTAACTGTTT |
| ZNF367 | CACCACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTAC<br>AAAAGTACTGCCCAGCATACTTAAAAAGTAGATCCTTGGGCAT<br>AAGCTAAGCACCTTATTTGCTTATCATAGGCTGCTATTCTGTAG<br>AAATTTATGAAGAATGTTATTGCCCCAGAATATGGGGTGAGAG<br>AGAACTGCACTTTTTTAATATGGAAATGAATTCATCGTAAAGT<br>TTAAAATATTTTGTAAATATGGACTGCACAGTACAGGGTAGAA<br>AACTACATATTGAACAACTCCGACAGTAGCTTGGACACTGACT<br>CTTCCACTGTACAAAAGTACTGCCCAGCATACTTAAAAAGTAG<br>ATCCTTGGGCATAAGCTAAGCACCTTATTTGCTTATCATAGGC<br>TGCTATTCTGTAGAAATTTATGAAGAATGTTATTGCCCCAGAA<br>TATGGGGTGAGAGAGAACTGCACTTTTTTAATATGGAAATGAA<br>TTCATCGTAAAGTTTAAAATATTTTGTAAATATGGACTGCACA<br>GTACAGGGTAGAAAACTACATATTGGGCTACTCCGACAGTAG<br>CTTGGACACTGACTCTTCCACTGTACAAAAGTACTGCCCAGCA<br>TACTTAAAAAGTAGATCCTTGGGCATAAGCTAAGCACCTTATT<br>TGCTTATCATAGGCTGCTATTCTGTAGAAATTTATGAAGAATG<br>TTATTGCCCCAGAATATGGGGTGAGAGAGAACTGCACTTTTTT<br>AATATGGAAATGAATTCATCGTAAAGTTTAAAATATTTTGTAA<br>ATATGGACTGCACAGTACAGGGTAGAAAACTACATATTGTCA<br>GACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTACAAA<br>AGTACTGCCCAGCATACTTAAAAAGTAGATCCTTGGGCATAAG<br>CTAAGCACCTTATTTGCTTATCATAGGCTGCTATTCTGTAGAAA<br>TTTATGAAGAATGTTATTGCCCCAGAATATGGGGTGAGAGAGA<br>ACTGCACTTTTTTAATATGGAAATGAATTCATCGTAAAGTTTA<br>AAATATTTTGTAAATATGGACTGCACAGTACAGGGTAGAAAA<br>CTACATATTGGTTT |

| | |
|---|---|
| mZNF367 | CACCACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTAC<br>AAAAGTACTGCCCAGCATACTTAAAAAGTAGATCCTTGGGCA<br>ACAAGACGTAGTGTTGACTCGCTAGATCTCATGTATATTTACA<br>TATTCTTTATGAAGAATGTTATTGCCCCAGAATATGGGGTGAG<br>AGAGAACTGCACTTTTTTAATATGGAAATGAATTCATCGTAAA<br>GTTTAAAATATTTTGTAAATATGGACTGCACAGTACAGGGTAG<br>AAAACTACATATTGAACAACTCCGACAGTAGCTTGGACACTGA<br>CTCTTCCACTGTACAAAAGTACTGCCCAGCATACTTAAAAAGT<br>AGATCCTTGGGCAACAAGACGTAGTGTTGACTCGCTAGATCTC<br>ATGTATATTTACATATTCTTTATGAAGAATGTTATTGCCCCAGA<br>ATATGGGGTGAGAGAGAACTGCACTTTTTTAATATGGAAATGA<br>ATTCATCGTAAAGTTTAAAATATTTTGTAAATATGGACTGCAC<br>AGTACAGGGTAGAAAACTACATATTGGGCTACTCCGACAGTA<br>GCTTGGACACTGACTCTTCCACTGTACAAAAGTACTGCCCAGC<br>ATACTTAAAAAGTAGATCCTTGGGCAACAAGACGTAGTGTTGA<br>CTCGCTAGATCTCATGTATATTTACATATTCTTTATGAAGAATG<br>TTATTGCCCCAGAATATGGGGTGAGAGAGAACTGCACTTTTTT<br>AATATGGAAATGAATTCATCGTAAAGTTTAAAATATTTTGTAA<br>ATATGGACTGCACAGTACAGGGTAGAAAACTACATATTGTCA<br>GACTCCGACAGTAGCTTGGACACTGACTCTTCCACTGTACAAA<br>AGTACTGCCCAGCATACTTAAAAAGTAGATCCTTGGGCAACAA<br>GACGTAGTGTTGACTCGCTAGATCTCATGTATATTTACATATTC<br>TTTATGAAGAATGTTATTGCCCCAGAATATGGGGTGAGAGAGA<br>ACTGCACTTTTTTAATATGGAAATGAATTCATCGTAAAGTTTA<br>AAATATTTTGTAAATATGGACTGCACAGTACAGGGTAGAAAA<br>CTACATATTGGTTT |
| hunchback | CACCTTGTTGTCGAAAATTGTACATAAGCCAAAACATTGTTGT<br>CGAAAATTGTACATAAGCCAAGGCTTTGTTGTCGAAAATTGTA<br>CATAAGCCAATCAGTTGTTGTCGAAAATTGTACATAAGCCAAG<br>TTT |
| mP_<br>hunchback | CACCTTGTTGTCGAAAATACAACATAAGCCAAAACATTGTTGT<br>CGAAAATACAACATAAGCCAAGGCTTTGTTGTCGAAAATACA<br>ACATAAGCCAATCAGTTGTTGTCGAAAATACAACATAAGCCA<br>AGTTT |

**Table C6: Primer sequences for generation of 4xLRIG3 miRNA seed controls.**
Controls were assembled as described in Li et al 2018.

| Name | Sequence |
|---|---|
| new4XLRIG_WT_M1_FP | GCGTCTCTCACCCCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACAAC AAG AGA CGC |
| new4XLRIG_WT_M2_FP | GCGTCTCTAACACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACGGC TAG AGA CGC |
| new4XLRIG_WT_M3_FP | GCGTCTCTGGCTCCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACTCA GAG AGA CGC |
| new4XLRIG_WT_M4_FP | GCGTCTCTTCAGCCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACGTT TAG AGA CGC |
| new4XLRIG_WT_M1_RP | GCGTCTCTTGTTGTTCACTTGAGGTAGTATGTTAAGCTTTTCCTTTGGGGTGAGAGACGC |
| new4XLRIG_WT_M2_RP | GCGTCTCTAGCCGTTCACTTGAGGTAGTATGTTAAGCTTTTCCTTTGGTGTTAGAGACGC |
| new4XLRIG_WT_M3_RP | GCGTCTCTCTGAGTTCACTTGAGGTAGTATGTTAAGCTTTTCCTTTGGAGCCAGAGACGC |
| new4XLRIG_WT_M4_RP | GCGTCTCTAAACGTTCACTTGAGGTAGTATGTTAAGCTTTTCCTTTGGCTGAAGAGACGC |
| new4XLRIG_mutM1_FP | GCGTCTCTCACCCCAAAGGAAAAGCTTAACATAGTGCGTAAAGTGAACAAC AAG AGA CGC |
| new4XLRIG_mutM2_FP | GCGTCTCTAACACCAAAGGAAAAGCTTAACATAGTGCGTAAAGTGAACGGC TAG AGA CGC |
| new4XLRIG_mutM3_FP | GCGTCTCTGGCTCCAAAGGAAAAGCTTAACATAGTGCGTAAAGTGAACTCA GAG AGA CGC |
| new4XLRIG_mutM4_FP | GCGTCTCTTCAGCCAAAGGAAAAGCTTAACATAGTGCGTAAAGTGAACGTT TAG AGA CGC |
| new4XLRIG_mut_M1_RP | GCGTCTCTTGTTGTTCACTTTACGCACTATGTTAAGCTTTTCCTTTGGGGTGAGAGACGC |
| new4XLRIG_mut_M2_RP | GCGTCTCTAGCCGTTCACTTTACGCACTATGTTAAGCTTTTCCTTTGGTGTTAGAGACGC |
| new4XLRIG_mut_M3_RP | GCGTCTCTCTGAGTTCACTTTACGCACTATGTTAAGCTTTTCCTTTGGAGCCAGAGACGC |
| new4XLRIG_mut_M4_RP | GCGTCTCTAAACGTTCACTTTACGCACTATGTTAAGCTTTTCCTTTGGCTGAAGAGACGC |