REGULATORY MECHANISMS AND RNA TARGETS OF HUMAN PUMILIO PROTEINS

by

Jamie Lynn Van Etten

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in the University of Michigan
2013

Doctoral Committee:

Assistant Professor Aaron C. Goldstrohm, Chair
Professor David Engelke
Professor Jairam Menon
Professor Alice Telesnitsky
Associate Professor David L. Turner

This work is dedicated to my dad.

Acknowledgements

It is important to note that this work happened with the help of many colleagues, professors, family members, and friends. It goes without saying that my mentor, Aaron Goldstrohm, deserves my sincere thanks. He provided me with technical training, intellectual guidance, and creative freedom to explore a multitude of scientific ideas.

I'd also like to thank Trista Schagat for setting the stage for this project in the year before I joined the lab, and for being a wonderful collaborator and mentor throughout my time in the lab: none of this work could have been done without her input and expertise. Thanks to my fellow graduate students in the Goldstrohm lab, Nate Blewett, Chas-E Weidmann, and Nathan Raynard. It was a pleasure learning from and with you, and I will always value their scientific ideas and expertise, collegiality, and friendship. I would also like to thank May Tsoi who not only brought joy to the lab and many smiles to my face every day, but whose wicked multi-tasking and technical skills helped me to finish projects without drowning in a sea of minipreps and LB broth.

Thanks to all the members of my committee, David Engelke, Alice Telesnitsky, Jerry Menon, and Dave Turner. Your insights, guidance, and tough questions were instrumental in the development and success of these projects. I'd also like to thank staff members of the Department of Biological Chemistry and program in Chemical Biology, specifically Beth Goodwin, Prasanna Baragi, Julie Woodroof, Amanda Howard, Wesley Tanner, Justine Altman,

Laura Howe, and Amy McGovern, for keeping me in line, caffeinated, and always well-fed with forbidden peanut butter cookies.

To my university professors who let me geek out in their presence and helped me to realize just how much I love science, Chris Korey, Pam Riggs-Gelasco, and Michelle Mac Brooks: this is all your fault. Kiz, even though I've moved to human tissues, I will [sadly] never forget how to virgin flies.

My friends, family, and cycling buddies (UM and AABTS!) deserve my thanks for providing me with a network of support, laughs, hammerfests, beer, and kind words throughout the last six years. Special thanks goes to Ed, Gabrielle, Tebo, and Beth: our many dinner parties, movie nights, emails, and all night conversations about science, life, and everything in between will be fond memories for many years to come. I would also like to thank Em, Nicole, Liz, Daniel, Shana, Josh, Mary, and Amy for your friendship, support, and for providing ice cream and baby therapy on occasion (Thank you, Edie, Oscar, and Jack!). Finally, I want to acknowledge my mom and sister for giving me your love and support, for feeding me and calling to make sure I'm alive after I've been in lab far too long, and for being present and letting me share with you the ups and downs of this process.

Table of Contents

List of Figures

List of Tables

Abstract

Gene expression requires balance between synthesis and decay of mRNAs. Aberrant expression of genes results in disease. PUF proteins are a family of eukaryotic RNA binding regulators characterized by a conserved RNA binding domain that binds 8-12 nucleotide sequence elements, called PREs, in target mRNAs. PUFs are thought to regulate processes including cell proliferation and memory formation. In model organisms, PUFs block translation and enhance deadenylation and mRNA decay. Humans have two PUFs: PUM1 and PUM2. Despite a substantial foundation of knowledge about PUFs in model organisms, the RNA targets of PUMs and mechanisms by which they repress remain unclear. I will present evidence herein that PUM1 and PUM2 are regulators in human cells. PUM1 and PUM2, which exhibit identical RNA binding specificities, employ a conserved mechanism of repression via direct recruitment of the CCR4-NOT deadenylase complex, causing translational inhibition and accelerated mRNA decay. PUMs also repress a non-adenylated mRNA; therefore, PUMs utilize a second, deadenylation-independent repression mechanism. RNAs regulated upon PUM knockdown in human cells were identified in this study and compared with RNAs in the genome that contain at least one PRE and previously classified, bound PUM mRNA targets to yield a dataset of 487 direct PUM targets. Direct PUM target mRNAs encode proteins involved in cancers, signaling, development, and neurological functions. This study demonstrates direct, potent PUM repression in human cells, illuminates mechanisms by which they enact repression, and identifies 487 direct RNA targets of PUM regulation; together revealing PUM regulated pathways in human cell networks.

CHAPTER 1

Introduction

Expression of genetic information encoded by eukaryotic genes involves numerous tightly controlled and coordinated processes. The central dogma asserts that DNA is transcribed to RNA, which is translated to protein (1). Among those steps exist intricate and dynamic regulatory networks that serve to ensure timely and proper gene expression. Each gene must be turned on or off at precisely the right time, at the right place, and in the right quantity; gene expression depends on the particular tissue, cell type, developmental stage, and biological response to a particular environmental stimulus. Aberrant gene expression results in disease.

DNA is transcribed to RNA by transcription. To produce a mature protein encoding messenger RNA, the RNA must be co-transcriptionally processed and exported from the nucleus to the cytoplasm. Throughout these processes, the mRNA is packaged with protein factors, together called messenger ribonucleoprotein particles (mRNPs), which ultimately must help the mRNA escape destruction by nuclear exoribonucleases and export the molecule to the cytoplasm (2). There, complexes of RNA binding factors, decay machinery, and the translational apparatus work to ultimately regulate gene expression; either by controlling translation or by interfering with mRNA stability and promoting its decay. Destructive enzymes and protein complexes seek

to destroy all but correct messenger RNAs and thus function as key quality control regulators in gene expression (3). Each of these steps in gene expression is subjected to a host of regulatory activities: transcriptional regulators positively and negatively affect synthesis of mRNA, regulators control splicing and alternative cleavage and polyadenylation to generate different forms of each gene (4, 5). Furthermore, mRNA localization, stability, and translation efficiency are subjected to stringent regulatory controls (6, 7). The work presented in this dissertation highlights the nexus of regulatory activity between expression of an mRNA transcript and its translation to protein using the PUF family of proteins as a model for regulation.

Messenger RNA processing and 3' end formation

Transcription of protein coding genes and many noncoding genes begins when RNA Polymerase II binds the promoter of the gene, together with many general transcription factors, to form a preinitiation complex (8). The DNA double helix is melted in an ATP dependent manner, and the template strand moves down into the base of the active site cleft of the polymerase (8). The polymerase adds several complementary ribonucleotides, pauses briefly, and then elongation of the transcript begins, during which sequential addition of nucleotides is processive, reaching rates of 1-4 kb/minute (9). During the early phase of transcription elongation; specifically within the first 40 nucleotides, a modified 7-methylguanosine cap is added to its 5' end (10, 11). During the elongation phase, introns are removed by splicing and, with the exception of histone mRNAs, mRNAs are subjected to cleavage and addition of a 3' poly (Adenosine) tail (10). Key structural elements including the 5' cap and poly (A) tail, as well as mechanisms of 3' end formation will be discussed subsequently.

The 5' 7-methylguanosine cap

The 7-methylguanosine cap is added to the 5' end and serves several purposes: it promotes splicing, mRNA export, and translation (12-17). Addition of the cap was confirmed by radioactive labeling experiments in both mouse myeloma and human HeLa cells (18, 19). In translation systems and wheat germ extracts, the presence of the 5' cap allows translation of mRNA to protein to proceed, whereas little or no translation is observed in its absence (17). The 5' cap plays a protective role by shielding the body of the mRNA from cellular 5'exonucleolytic degradation (20). In addition, the cap promotes translation (21) Poly (A) binding protein (PABP) is bound to the poly (A) tail at the 3' end, and interacts with cap-bound eIF4G to circularize the mRNA: it is thought that translation is facilitated by formation of this closed loop structure (22).

The poly (Adenosine) tail

Termination of transcription of Pol II transcripts, cleavage, and polyadenylation are intimately linked. The processing machinery in mammals that carries out mRNA 3' end formation is complex, consisting of upwards of 80 different proteins (23-25). The length of the poly (A) tail varies widely among organisms. For example, in yeast the tail is approximately 80 adenosines, whereas in mammals the poly (A) tail is between 200 -250 adenosines in length (26).

The poly (A) tail serves many functions: it protects the 3' end of the mRNA from degradation and promotes translation of the coding region by binding to regulatory factors. The poly (A) tail is formed by cleavage and polyadenylation, often co-transcriptionally. Within the mRNA, a number of sequence elements direct this process: the poly (A) signal is located approximately 20 nucleotides upstream from the poly (A) site, where the poly (A) tail is added, and a downstream U/GU rich element further direct the process (5, 27, 28). Upstream and downstream auxiliary sites affect cleavage and polyadenylation as well. While a canonical

AAUAAA poly (A) signal is often found in mRNAs, several noncanonical alternative poly (A) signals have been identified and are thought to make up ~ 14% of poly (A) signals in humans and mice (5). The strength of the poly (A) site affects the efficiency of transcriptional termination: weakened poly (A) sites in human β-globin and α-globin genes reduce termination frequency and efficiency (29-31).

Canonical cleavage and polyadenylation occurs at the end of the elongation phase of transcription when cleavage and polyadenylation specificity factor (CPSF) interacts with the carboxyl terminal domain (CTD) of RNA Pol II: it recognizes the poly (A) signal, AAUAAA or another non-canonical signal, and then at the downstream GU-rich processing signal the protein CStF (cleavage stimulatory factor) binds and recruits CPSF, by which Pol II is released, and the transcript is cleaved and polyadenylated (9). After co-transcriptional cleavage and polyadenylation, the newly generated RNA fragment has an unprotected 5' end and is left vulnerable to exoribonucleolytic degradation by XRN2 (9).

In addition to cleavage and polyadenylation, mRNAs are alternatively polyadenylated to generate different 3' untranslated regions (5). More than half of mRNAs are alternatively processed: this depends on cell type, disease state, and cell cycle and differentiation (32, 33). Alteration of the mRNA 3' end has a number of implications for localization, stability, function, and regulation of the mRNA (5, 32, 34). For example, mRNAs are alternatively polyadenylated after upstream poly (A) signals to generate shortened 3'UTRs upon activation of primary T lymphocytes in mice (32). Further, alternatively polyadenylated, shortened 3'UTRs correlate with cell proliferation (32). Thus, changing the 3' UTR of an mRNA results in significant changes to its expression and regulation by RNA binding proteins, its translation, as well as its response to environmental changes.

Poly (A) binding proteins in mammals bind and coat the poly (A) tail and play roles in translation and mRNA stability. Nuclear PABPs exist in mammals, and function to stimulate polyadenylation of the mature mRNA transcript (35, 36). Poly (A) tails in the cytoplasm are coated with multiple poly (A) binding proteins (PABP), which bind specifically to poly (A) through four RNA recognition motifs (RRMs) in their N terminal regions (35). Mammals have 5 canonical cytoplasmic PABPs, though the most widely studied PABP in mammals is PABC1, which is the major form of the protein (36). PABC1 has several functions: the 70 kDa protein coats the tail and interacts with the cap-binding complex, bringing the mRNA into a closed loop structure, thereby stimulating translation (35). PABC1 is ejected from the poly (A) tail as the tail is shortened by deadenylases, but may also play roles in recruitment of decay factors (37).

Translation and mRNA decay are interconnected processes

Mature mRNA encodes protein; however, less than half of its sequence contains the open reading frame (ORF) and the remainder consists of regulatory sequences (38). The ORF is flanked on either side by 5' and 3' untranslated regions: sequence that is not translated to protein, but instead specifies regulatory instructions for the gene product. Many sequence elements located in the body of the mRNA interact with RNA binding proteins and complexes to positively or negatively affect translation and stability of the mRNA. Examples of specific regulatory activities conferred by *cis* regulatory sequences and *trans* acting factors will be discussed in later sections.

After co-transcriptional processing and 3' end formation, the mature mRNA is exported from the nucleus to the cytoplasm in a process that is ATP-dependent (39, 40). There in the cytoplasm, the mRNA can engage with the translational apparatus or decay machinery. Upon export, poly (A) binding protein (PABP), which binds with high affinity to short poly (A)

stretches, binds and coats the poly (A) tail (41). Together, the 5' cap and poly (A) tail protect and ensure the stability of the messenger RNA and synergize to promote translation of the mRNA (42, 43).

Translation and decay are opposed in function; therefore, each process is tuned to promote proper spatial and temporal control of mRNA degradation and protein synthesis. Initiation of translation is rate limiting and occurs when the 40S – mRNP complex and 60S ribosomal subunits join in a highly ordered fashion. The nascent protein is then produced while the mRNA is held in a circularized conformation, joined together by cap bound initiation factors and poly (A)-bound PABP. In this conformation, the mRNA can load multiple ribosomes, to form polysomes, along its coding region and be translated efficiently many times: it is estimated that translation rates reach 750-1300 proteins per mRNA per hour (44). Disruption of the closed loop structure and dissociation of translation factors is thought to make the mRNA accessible to decay factor binding, which allows degradation to proceed rapidly. Thus, translation and decay are intricately linked processes and work in opposition to control gene expression; these processes and their interrelationship will be described in the ensuing sections.

Regulation of translation initiation

Translation begins with initiation, a process that is rate limiting and extensively regulated. The mRNA is activated when an mRNP complex consisting of mRNA bound to eIF4A, eIF4B, eIF4E, and eIF4G and PABP. eIF4E binds directly to the 5' cap and PABP coats the poly (A) tail (45). During initiation, a ternary complex consisting of eIF2 and the initiator methionine tRNA forms (45). Initiation factors eIF1, eIF1A, eIF3, and eIF5 aid to recruit the ternary complex to the 40S ribosomal subunit to form the preinitiation complex (PIC) (45). The activated mRNP is recruited to the preinitiation complex, at which point the PIC binds the 5' end

of the mRNA and begins to scan along the molecule to find the first methionine codon, called the start codon (AUG) (46-48). The sequences flanking the AUG start codon are crucial for enhancing translation of mRNAs; called the Kozak sequence, the consensus motif (GCC)GCC(A/G)CCATGG promotes efficient translation (46, 49).

Once the start codon is recognized, eIF1 can be released and GTP on eIF2 is hydrolyzed to GDP, a process that stops scanning until the 60S subunit joins. At this point, eIF2-GDP and eIF5 are released so that eIF5B can aid in joining the 60S subunit to the 40S subunit bound to the mRNP and remaining initiation factors (45). Once the two subunits have united, another molecule of GTP is hydrolyzed and initiation factors are released, allowing translation of the open reading frame to protein (45).

mRNAs are subjected to a very broad range of translation efficiencies (44). Initiation factor activity is often regulated by phosphorylation. For example, phosphorylation of eIF-2α inhibits translation, by blocking its recycling by the eIF2B, a GDP/GTP exchange factor (50). Conversely, phosphorylation of eIF-4E appears to enhance its cap-binding activity and thereby increases translation (50). The cap promotes translation by binding eIF4E, which recruits eIF4A and eIF4G to form the eIF4F complex to ultimately recruit the 43S pre-initiation complex and form the 48S initiation complex (21, 22). In addition to regulation of components of the translational machinery, elements in the mRNA affect translational efficiency as well. Secondary structure in the 5' UTR inhibits translation efficiency (51). RBPs can bind elements in the 3' UTR to regulate translation, either by activation or repression (52-54). One of our goals in this work and beyond is to determine if and how human PUF proteins, PUM1 and PUM2, affect translation.

mRNA deadenylation and decay

One consequence of gene expression is that mRNA will be degraded: its destruction is necessary to turn off genes. Transcription rates and mRNA decay work in opposition to dictate steady state levels of mRNAs. mRNA levels and half lives vary over several orders of magnitude, and range from minutes to many hours (44). In mammals, the median half life is 7-10 hours, and ranges from <1 hour to several hundred hours, though these values depend largely on cell type and environmental conditions (44, 55-57). mRNAs encoding structural or housekeeping genes such as GAPDH, for example, have very long half-lives; conversely, mRNAs encoding signaling genes have very short half-lives, making them available only briefly for translation. A genome-wide study predicted rates of synthesis of proteins and mRNAs: certain groups of genes, arranged by function, evolved characteristic half lives (44). For example, genes with unstable mRNAs and proteins are involved in strictly timed and regulated processes such as mitosis and regulation of transcription; whereas stable mRNAs and their stable protein transcripts are involved in housekeeping processes such as metabolism and translation (44). All this is to say that mRNA decay must be carefully controlled such that expression of the appropriate amounts of proteins are properly timed and spatially controlled: without high levels of regulation, aberrant gene expression will occur and disease will likely follow. A range of regulatory mechanisms is therefore necessary to control mRNA abundance and degradation.

Deadenylation initiates mRNA decay

Deadenylation, or 3' exonucleolytic digestion and removal of the poly (A) tail, is often the rate limiting step of mRNA decay (41, 58, 59). Shortening of the poly (A) tail acts as a dimmer switch: longer tails promote translation and shortened tails act as a signal to RNA decay factors that will go on to degrade the mRNA (Fig. 1.1). Pathways by which mRNAs are further

degraded occur by either decapping and 5' degradation or 3' degradation (Fig. 1.1). Over the lifetime of specific mRNAs, the poly (A) tail may change in length depending on the phase of the cell cycle or environmental conditions: for example, during oocyte development (41, 60).

There is diversity among deadenylase enzymes in eukaryotes. All known deadenylases belong to one of two superfamilies; EEP (exo-nuclease-endonuclease-phosphatase) or DEDD type deadenylases, named for their active site residues (41). DEDD type deadenylases employ catalytic aspartate and glutamate residues dispersed between three Mg^{2+}-coordinating exonuclease motifs and include POP2 orthologs, PAN2, and PARN families (41, 61). EEP type enzymes are also Mg^{2+}- dependent and include CCR4 orthologs: these enzymes contain conserved catalytic aspartate and histidine residues in their putative active sites (41, 62).

Multiple deadenylases can work together to degrade poly (A) tail: CCR4 and POP2 deadenylases work together as a heterodimer, yet they can also work sequentially as in the case of PAN2 and the CCR4-POP2 heterodimer. In mammals, the enzyme PAN2 in complex with PAN3 first shortens the poly (A) tail to approximately 110 nucleotides (63). In yeast, this complex regulates shortening of the poly (A) tail in the nucleus, whereas CCR4 is the predominant deadenylase in the cytoplasm (41, 64). After initial shortening of the poly (A) tail in mammals, degradation of mRNA occurs by decapping and 5' decay or by further deadenylation and decay from the 3' end (Fig. 1.1). The remainder of the poly (A) tail is removed by CCR4 and POP2 deadenylase enzymes (63). Ccr4p and Pop2p are required for cytoplasmic deadenylation in yeast; if Ccr4p is overexpressed in a *Δpop2* strain, defects in deadenylation are rescued suggesting that Ccr4p is the major deadenylase (65, 66). After deadenylation by the CCR4-NOT complex, the remaining mRNA body can be degraded by the exosome in the 3'-5' direction (3, 67, 68).

CCR4 and POP2 are components of the CCR4-NOT deadenylase complex, which exists in several forms in mammals depending on its subunit composition (69, 70). This heterogeneous complex is formed by core non-catalytic components, including CNOT1, CNOT2, CNOT3, CNOT9, CNOT10 (71). CNOT4, which in yeast is called Not4, is not stably associated with core components of the complex as it is in yeast (71, 72). The CCR4-NOT complex also contains two catalytic subunits, orthologs of CCR4 and POP2. Mammals have two active CCR4 orthologs, CNOT6 and CNOT6L, as well as two POP2 orthologs, CNOT7 and CNOT8 (71, 73-77). These orthologs can form heterodimers, contributing to variation within CCR4-NOT complexes. Therefore, heterodimerization of CCR4 and POP2 orthologs results in the formation at least 4 different multisubunit CCR4-NOT complexes in mammals.

Decapping factors cause removal of the 5' cap and potentiate rapid mRNA decay

Decapping normally occurs after shortening of the poly (A) tail, at which point the 5' cap will be removed by decapping enzymes and the remainder of the mRNA will be degraded by XRN1, which is the major cytoplasmic exoribonuclease (Fig. 1.1)

(78-80). The 5' cap structure is typically removed by a heterodimer of DCP1 and DCP2 in yeast, whereas in humans the decapping complex consists of DCP1 and DCP2 linked via HEDLS protein (3). In mammals, upwards of seven different decapping enzymes have been identified including DCP2 and NUDT16, as well as several other NUDT family members (81). Decapping in yeast requires removal of the cap-bound eIF4E (82). Several factors enhance decapping and promote 5'-3' mRNA decay by the exosome, including the LSM1-7-Pat1 complex, which helps to coordinate decapping after deadenylation by binding the short poly (A) tail of deadenylated mRNA (80, 83).

Interplay between translation and mRNA decay

During translation, the mRNA is held in a closed loop by interactions with RBPs at the 5' and 3' ends (43, 84). Displacement of initiation factors and RBPs linking the 3' or 5' ends must occur before decay factors gain access to the mRNA. Ccr4p deadenylase activity is inhibited by excess Pab1p, which coats the tail to promote translation(66). Shortening of the poly (A) tail, a process which ejects PABP and causes the disruption of the closed loop structure, initiates the decay process and thereby inhibits translation, as poly (A) binds initiation factors to enhance translation (58, 78, 85-88).

Furthermore, blocking translation elongation in yeast with cycloheximide stabilizes the MFA2 mRNA, suggesting that modulating translation has a direct or indirect effect on decay (15). In yeast, bound eIF4E blocks decapping activity in yeast by interfering with Dcp1 binding to the cap; once eIF4E dissociates, the decapping complex is free to bind and initiate the process of mRNA decay (82). Mutations that inactivate translation initiation factors result in an increased rate of mRNA decay, which illustrates that regulatory proteins affecting one process directly can have secondary effects on other steps (89). I will discuss examples of such interplay throughout this work. PUF proteins enhance deadenylation, which can disrupt translation and cause mRNA decay. Alternatively, PUFs may block translation directly: as a consequence, decay is enhanced. Identifying direct interactions between PUFs and deadenylases is therefore crucial for discrimination between these models.

3' UTR *cis* elements and *trans*-acting factors confer post-transcriptional regulation

It is likely that hundreds of RNA binding proteins exert combinatorial control of gene expression (90-93). At the foundation of these processes lie direct interactions between mRNA and RNA binding factors, including proteins and other RNAs. Sequences embedded within the

mRNA transcript direct a cadre of regulatory complexes to positively or negatively affect its translation and decay. RNA binding factors act in *trans* to recognize these elements to activate or repress gene expression. Many of these RNA binding factors and their target sequences have yet to be characterized, though one thing is known: RNA binding factors are responsible for immense post-transcriptional regulation and precision and loss of their ability to control gene expression can cause catastrophic problems in biological processes (94).

Mutations and changes in expression of RBPs results in a number of diseases: for example, increased expression of the RBP eIF4E is found in transformed cancer cells (94). Additionally, defects in microRNA function or expression is associated with many cancers (95, 96). Aberrant changes in mRNA sequences can have drastic physiological consequences. For example, repeats of 70-120 CGG triplets in the 5'UTR of the FMR1 gene results in increased FMR1 mRNA levels and are associated with the Fragile-X Tremor Ataxia Syndrome (FXTAS), which causes late onset disruption of neurological function (97). Similarly, CAG expansions result in Huntington's disease and spinocerebellar ataxia (SCA) and CUG repeat expansions resulting in a gain-of-function mutant mRNA are associated with myotonic dystrophy (97). With the discovery of hundreds of new RNA binding proteins, it is clear that a great deal of information remains to be discovered regarding interactions between mRNA sequences and RNA binding factors, and how they affect normal physiology and disease.

Posttranscriptional control of gene expression depends on *cis* elements and *trans* acting factors

Cis elements are sequences found in the body of the mRNA, most frequently in the untranslated regions. Broadly, they provide a roadmap for regulation of the gene in which they reside: that is, they direct regulatory complexes to the body of the mRNA so as to control its

expression in different ways. C*is* elements direct localization of the mRNA in the cell, activate or repress protein expression, and stabilize or destabilize the mRNA (60, 98-100). They integrate cellular signals and act as tunable regulators to adjust levels of protein expression and to control mRNA stability and decay. These sequences recruit *trans* acting factors, which include translational machinery and decay factors. Several well-characterized examples of *cis* elements and associated *trans*-acting RNA binding factors including miRNA response elements, AU rich elements (AREs), and Pumilio response elements (PREs) will be discussed in subsequent sections, which will focus on highlighting known mechanisms and trends in their regulatory activities.

miRNAs are *trans*-acting factors that inhibit mRNA translation and promote decay

In the early 1990s, the Ambros and Ruvkun labs made striking discoveries related to a gene, *lin-4*, that appeared to be involved in timing cell division between the first and second larval stages in *C. elegans*. The small *lin-4* RNA, only 22 nucleotides in length, had antisense complementarity to a region in the 3' UTR of the *lin-14* mRNA, and somehow directed translational repression of its protein product, LIN-14 (101, 102). Over the ensuing years, many hundreds of small RNAs, called microRNAs (miRNAs), were discovered to bind to sites in the 3' UTR of as many mRNAs to down-regulate their expression (103).

MicroRNAs are encoded in the genome as long transcripts called primary miRNAs (104). These transcripts are processed twice: first, the pri-miRNA is cleaved by the RNase III endonuclease Drosha, which is specific for dsRNA, to a smaller 60-70 nucleotide stem loop structure called a precursor miRNA, or pre-miRNA (105). The shortened pre-miRNA is transported from the nucleus to the cytoplasm, where it subjected to final processing steps.

Transport of the pre-miRNA requires the karyopherin family member Exportin-5, a nucleo-cytoplasmic nuclear export factor that uses the GTPase Ran-GTP (106, 107).

Once localized to the cytoplasm, the pre-miRNA is cleaved by another RNase III endonuclease, Dicer, to its final 22 nucleotide structure (108-110). The final dsRNA miRNA product is then deployed to bind and repress its targets along with a mRNP complex called miRISC consisting of a conserved Argonaute (AGO) proteins and GW182 (in humans, its orthologs are TNRCA1, 2, and 3) (111). MiRNAs basepair with perfect or imperfect complementarity to miRNA response elements in the 3'UTR of target mRNAs, which indicates that a single miRNA can bind and regulate a number of different mRNAs (103, 111-114).

MicroRNAs in complex with the miRISC machinery bind a target sequence in an mRNA sequence and recruit regulatory factors to it to inhibit its expression. miRNAs alter gene expression of over half of mammalian genes in a number of ways: they cause mRNA decay by recruitment of deadenylases, and inhibit translation before and after initiation (112). miRNAs inhibit cap-dependent translation at the initiation step (115-119). Another proposed mechanism of miRNA mediated translation inhibition involves blocking translation at steps beyond initiation (120-123). A host of studies implicate miRISC in enhancing mRNA decay by interacting with decapping factors and deadenylation enzymes (124-126). GW182-mediated repression by miRISC correlates with its ability to bind PABP, CCR4-NOT deadenylase complex member NOT1, and the deadenylase PAN2 (127, 128). Intriguingly, a new study showed that the miRNA-associated scaffold protein GW182 can also reduce association of PABP and eIF4E with the poly (A) tail and 5' cap in order to block translation independently of deadenylation in *Drosophila* (129). In humans, GW182 has three paralogs, TNRC6A, B, and C (129). Many of

the protein corepressors that bind miRISC are recruited by other *cis* elements, which will be described in subsequent sections.

AU-rich elements positively and negatively affect mRNA translation and stability

One of the first examples of a *cis* element responsible for conferring gene regulation is the AU rich element (ARE) (130). AREs are found in the 3' UTRs of many genes including chemokines, cytokines, and proinflammatory enzymes (131). Early experiments suggested that the sequence AUUUA in the 3'UTR of mRNAs triggers the rapid and selective degradation of the transcript(130). Typically this ARE is embedded within a U-rich setting (131). However, over the years it was discovered that there are at least three classes of AREs within 3'UTRs: class I contains several copies of the AUUUA sequence within a U-rich environment, class II contains two overlapping UUAUUUA(U/A)(U/A) nonomers, and class II contains regions rich in Us but do not contain the canonical AUUUA sequence (132). Multiple ARE binding proteins (ARE-BPs) have been identified and are able to bind to many mRNAs, often with no preference for a particular class of ARE (132). ARE-BPs recruit deadenylase and decapping factors, as well as the exosome to repress ARE mRNAs (133-135).

AREs function as bifunctional switches: they direct activation or repression of a specific mRNA by recruiting different RNA binding proteins. An early example of regulation of an ARE mRNA by an ARE-BP was demonstrated using a β-globin reporter. A highly conserved AT-rich sequence found in the 3' UTR of the gene coding for lymphokine granulocyte-monocyte colony stimulating factor (GM-CSF) was inserted into the 3'UTR of rabbit β-globin gene and caused significant reduction of globin transcripts and decreased mRNA stability (130). Deletion of an AU rich element from the c-*fos* mRNA results in stabilization of the gene and slowing of deadenylation (85).

Tristetraprolin (TTP) is a negative regulator that binds class II AREs elements through its tandem zinc finger domain and promotes decay of the mRNA target (132, 136). TTP functions in physiological processes as a tumor suppressor as well as in inflammatory processes (136). The ARE-BP TTP binds with very high affinity to the ARE containing IL-2 mRNA: in TTP knockout mice, IL-2 protein is overproduced and its mRNA is stabilized (137). It was later demonstrated that TTP associates with decapping enzymes and CNOT1, a subunit of the CCR4-NOT deadenylase complex to repress its targets (135, 136). By recruiting deadenylases and mRNA decay factors, TTP acts as a potent repressor of target mRNA expression.

As discussed previously, AREs are bifunctional: the ARE-BP HuR, a member of the RRM superfamily, is a stabilizer of bound ARE mRNAs. One example of HuR stabilization is the regulation of the VEGF mRNA. Under hypoxic conditions, the vascular endothelial growth factor (VEGF) mRNA undergoes significant stabilization and, under these conditions, mediates angiogenesis, which has important implications in cancer and heart disease (138). HuR stabilizes of the VEGF mRNA by binding to its ARE under hypoxic conditions (138). There are many potential targets of HuR activity, which were identified using PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation). (139, 140). PAR-CLIP is a technique during which photoreactive ribonucleoside analogs are incorporated into RNA transcripts in live cells (141). Labeled transcripts are then crosslinked to bound proteins (RBPs), and immunoprecipitated from lysates (141). RNAs bound to immunoprecipitated proteins are isolated and converted to a cDNA library, which can be sequenced by high throughput techniques (141). The mechanism(s) by which HuR stabilizes ARE mRNAs is (are) not fully understood, though recent transcriptome-wide studies suggest that HuR is involved in pre-

mRNA processing and mature mRNA stability (139, 140). Alternatively, HuR may work by displacing repressors by competitively binding to the ARE.

PUF response elements recruit PUF proteins to cause mRNA decay

The PUF response element (PRE) is a *cis* element that directs regulation of mRNAs by binding to PUF proteins and enhancing mRNA decay. PUF proteins, named for the founding family members Pumilio in *Drosophila melanogaster* and FBF in *C. elegans*, are conserved eukaryotic RNA binding proteins characterized by a homologous RNA binding domain (RBD) (142-146). A PRE is an 8-12 nucleotide consensus sequence, typically in the 3'UTR of target mRNAs. PUFs regulate genes underlying many important biological processes in model organisms, including stem cell proliferation, development, and neuronal function (143, 145-163).

The proposed ancestral function of PUFs is that they support germline stem cell maintenance, though they are involved in many processes (151, 164). For example, during *Drosophila* development, maternal hunchback (hb) mRNA is repressed at the posterior of the developing embryo by Pumilio (PUM) and Nanos (NOS). Here, the Pumilio RNA binding domain binds to the PRE in the 3' UTR of *hb* originally called the Nanos Response Element (NRE) with low nanomolar affinity (142-144, 165).

While PUFs in model organisms repress target mRNAs, the mechanisms by which they do so in humans remain unclear. PUFs in model organisms are thought to employ a number of distinct mechanisms to repress translation of specific mRNAs: they recruit deadenylation and decay machinery to destabilize the mRNA, and they directly inhibit translation (145, 160, 166-174).

PUF proteins repress messenger RNAs

There exist many potential targets of PUF regulation, and PUFs form complexes with many different regulatory proteins to carry out repression. PUFs have no known catalytic activity themselves, and thus are thought to serve as a scaffold that binds the mRNA PRE and regulatory complexes, to then direct regulatory complexes to the mRNA to carry out repression by interfering with translation and promoting mRNA decay. For example, yeast PUFs bring the Ccr4p-Pop2p deadenylase complex in proximity to the poly (A) tail of the mRNA, such that the Ccr4p deadenylase subunit may shorten the tail, signaling to decay complexes that the RNA is to be degraded.

In addition to causing repression of mRNAs, there is some evidence that suggests a role for PUFs in mRNA localization (175, 176). Yeast Puf3p was reported to localize to the mitochondria, where it localizes mRNAs and is thought to play a role in mitochondrial biogenesis (176, 177). Yeast Puf6p represses translation of the *ASH1* mRNA in the bud tip before the mRNA is localized to the bud cortex during late anaphase (178). Limited evidence also suggests a role for PUFs in activation of RNAs (179). It is unclear whether PUF activation is widespread, and this intriguing role of PUFs should be extensively tested.

PUFs bind RNA via a conserved binding domain

PUFs bind RNA nucleotides via a conserved modular RNA binding domain, consisting of 8 imperfect alpha-helical repeats of 36 amino acids each (142, 145, 146, 180-182). Each RNA base makes contacts with an individual repeat, with binding conferred by hydrogen bonding, van der Waals contacts, and base stacking interactions with exposed RNA recognition amino acid side chains along the inner, concave surface of the RBD (Fig. 1.2) (181). The outer surface of the crescent may contact co-repressors. Despite similarities in their modular structure, PUF RNA

binding domains vary among species in terms of their specificity of nucleotide binding. They bind 8-12 nucleotide RNA sequences, which differ with the exception of the first three ribonucleotides, UGU, which are required for binding (142). Such modularity welcomes efforts to alter binding specificity of PUFs, which has tremendous implications for research and clinical applications: for example, PUFs have been re-engineered to bind different PRE sequences such that in the future, it is feasible that they could be used to bind new RNA targets selectively (183-186).

PUF regulatory mechanisms in mammals have yet to be determined. Mammals have two PUF proteins, PUM1 and PUM2, which are orthologous to canonical PUF family members in other organisms. Human PUM1 and PUM2, which are the focus of this study, share 83% similarity and 75% identity, respectively (187). Their 361 amino acid long PUM RNA binding domains, or RBDs, share 91% identity and 97% similarity, and they differ in sequence primarily in their amino termini (187). PUM1 and PUM2 bind the consensus sequence UGUANAUA where N is any nucleotide (188, 189). Outside the conserved RNA binding domain, the amino and carboxyl terminal extensions of PUFs across species are divergent outside of their highly conserved RNA binding domains.

Global analyses of PUF gene regulation

Efforts to characterize the breadth of PUF regulation in model organisms have been made in recent years. RIP-Chip, RNA immunoprecipitation and microarray, is a method often used to identify bound RNA targets of proteins (190). Targets of all five yeast PUFs were identified by RIP-Chip assays in 2004 and thus provided new testing ground for understanding mechanisms of PUF regulation (191). Global analysis of *Drosophila* PUM has also been performed (191, 192). Together, studies in model organisms illustrate that PUFs bind hundreds of potential target

RNAs. Mammalian PUM1 and PUM2 bind and potentially repress hundreds of genes between them, as indicated by RIP-Chip (161, 163, 188, 189, 193, 194). In 2008, RIP-Chip studies indicated that purified PUM1 and PUM2 RNP complexes from human HeLa cells bound to 1424 and 751 RNA transcripts, respectively (188). Perhaps most striking was the finding that a high percentage, 88%, of PUM2 targets overlapped with those of PUM1 (188). Another RIP-Chip study performed in HeLa cells looked at bound targets of PUM1 and reported that 726 genes were bound to PUM1 (189). The authors used qRT-PCR to assess stabilization of a handful of target mRNAs upon PUM1 knockdown. Targets that were modestly stabilized upon PUM depletion include PCNA, Cks2, Cyclin B1, Cyclin E2, SLBP, and PUM2. RIP-Chip analysis in mice to identified 1527 putative PUM1 target genes (161).

Deletion of PUM1 in mice caused no apparent defects in lifespan, though they are smaller than wildtype mice and males exhibit testicular hypoplasia (161). This study was intriguing, in that it identified repression of many upstream targets of the p53 tumor suppressor gene by PUM1 in its analysis, shedding light on a potential regulatory role of PUMs in cell proliferation and disease (161). PAR-CLIP was used to identify targets of PUM2 upon overexpression in HEK293 cells and obtained ~3000 genes associated with PUM2 (141).

It is important to note that limited functional validation was performed in the above referenced studies. Less than half of the RIP-Chip and PAR-CLIP targets had identifiable PREs. Thus, it is likely that many targets identified were indirectly associated with PUMs. Bound RNAs were identified, though only in a few cases was mRNA stability of PUM targets assessed. Furthermore, an important consideration of these studies is that, in the majority of cases, mRNAs bound to only one PUM protein were identified. In addition to the observations described above that PUM1 and PUM2 targets significantly overlap, omitting one or the other from global target

analysis will inevitably weaken and cloud interpretation of the data (188). While RIP- ChIP and PAR-CLIP experiments provide valuable insight, functional analysis of regulation by PUMs is necessary to fully define their target mRNAs.

PUFs enhance deadenylation and mRNA decay

PUF regulation in model organisms correlates with mRNA decay. In yeast, PUFs enhance deadenylation of target mRNAs through a direct interaction with the CCR4-POP2 deadenylase complex (173, 195). This direct interaction between PUF and POP2 is conserved in yeast, *C. elegans,* and *D. melanogaster*. (160, 166, 167, 196, 197).

Yeast PUFs promote the decay of target mRNAs by recruitment of deadenylases and decapping factors (173, 174, 195, 198, 199). In yeast, PUFs bind directly to the Pop2p deadenylase subunit and thereby bring the remainder of the CCR4-POP2 deadenylase complex in proximity to the target, allowing Ccr4p to carry out deadenylation of the mRNA (173, 195, 199). Based on the high degree of conservation among RBDs and interaction with deadenylases, it begs the question: do human PUMs regulate mRNA via deadenylation? That question will be explored in Chapter 3 of this thesis, within the broader context of investigating specific mechanisms of human PUM repression.

Evidence from yeast also indicates that PUFs can enhance decapping of target mRNAs. An intriguing new mechanism of yeast Puf5p regulation involves the eIF4E binding protein Eap1p and its association with decapping factors to cause repression. Eap1p is required for Puf5p repression and is proposed to associate with Puf5p and Dhh1p to accelerate decapping of the HO mRNA (174). Conceptually, it is then feasible that human PUMs could affect mRNA decapping or other decay steps including exonucleolytic degradation following deadenylation and decapping.

One of the founder PUF family members, *Drosophila* Pumilio, represses translation of the hunchback mRNA to control anterior-posterior formation in the developing embryo. The development of *Drosophila* abdominal segments involves localization of two factors, *bicoid* mRNA at the anterior and *nanos* mRNA at the pole plasm: translation of each gives rise to gradients originating at each pole of the embryo (144). Maternally deposited Pumilio and Nanos repress *hunchback* (*hb*) mRNA in the developing embryo before the zygotic genome is expressed after the maternal to zygotic transition (143, 144).

Pumilio is maternally deposited and ubiquitously expressed throughout the embryo. Early genetic experiments found that cold sensitive *Drosophila pum* mutant embryos have defects in abdominal segmentation: only two of eight segments were formed and later experiments showed that Pumilio is essential for formation of abdominal segments (150, 153). Mutations in *pum* and *nos* cause defects in embryos due to absence of *hb* repression (153, 200). It was later shown that Pumilio could recognize and bind the NRE (144). It was originally surmised that Pumilio represses *hb* by activation of deadenylation, based on the correlation between repression and poly (A) tail shortening, which was thought to cause later translational inhibition(160). Years later, the Lehmann laboratory injected synthetic mRNAs lacking poly (A) tails or non-adenylated histone stem loop RNAs which remain unadenylated into *Drosophila* embryos and demonstrated that a poly (A) tail is unnecessary for Pumilio repression (167). However, the efficiency of repression of the non-adenylated reporter was reduced, which argues that the tail plays a role in PUM repression (167). Together these data support a deadenylation dependent mechanism of PUM regulation in *Drosophila.*

PUFs repress translation

Repression of RNAs by PUF proteins is correlated with reduced translation; this mechanism could be a direct repression of translation or could occur via secondary effects of mRNA deadenylation and decay. Limited data is available that demonstrates a direct effect of PUFs on translation. *In vitro* cell-free translation assays were used to measure direct effects of PUFs on translation: the authors found that recombinant yeast and *C. elegans* PUFs repressed reporter mRNAs that contain 3' UTR sequences of four known mRNA targets (169). While it is known that yeast Puf5p interacts with the CCR4-POP2 deadenylation complex to repress the HO mRNA, the Wickens lab sought to characterize translational repression of new RNA targets using *in vitro* biochemical assays (169). Yeast Puf5p was added to *in vitro* translation extracts along with reporter RNA containing the 3'UTRs of target genes; then, using this assay, the authors demonstrated translational repression of *HO*, *CIN8*, *DHH1*, and *RAX2* reporter RNAs (169). Another study implicates PUM2 from *Xenopus* in translational control by directly interfering with 5' cap recognition, thereby blocking initiation complex assembly on the cap (171).

In a recent study, *in vitro* experiments demonstrated that a complex consisting of human PUM2, Ago, and eIF1A formed to block translation elongation independently of miRNA association (170). The authors report that PUM2 binds Ago and eEF1A, which is a GTPase required for elongation: this complex is purported to inhibit GTPase activity and block translation.

PUMs in physiological processes

Evidence in the literature suggests that mammalian PUMs are involved in many biological processes, including stem cells, cell proliferation, spermatogenesis, and neural

function (141, 159, 161, 163, 188, 189, 193, 194, 201). We know that they bind many hundreds of targets from RIP-Chip and PAR-CLIP studies, but very little evidence ties human PUM binding and regulation to physiological function (141, 161, 188, 189). At the time we began this work, human PUM regulatory activity was relatively unstudied in functional assays: by association with their relationship to yeast PUFs, mammalian PUMs were believed to be repressors but rigorous examination of direct PUM1 and PUM2 repression was absent from the literature. In Chapter 3, I will discuss our finding that human PUM1 and PUM2 are potent repressors on their own and each can compensate for loss of its paralog. Many studies have been done to attempt to clarify the roles of human PUM1 and PUM2 in distinct biological phenomena and have provided insight into their functions in mammals. A great deal of information pertaining to localization and expression of PUM1 and PUM2 is available in the literature and will be discussed in this section. Furthermore, recent literature that proposes putative mechanisms of PUM repression in humans will be discussed through examples of mammalian systems that study the physiological functions and demonstrated targets and binding partners of human PUM proteins, thereby illustrating and highlighting their significance and the need for further experimental inquiry.

Mammalian PUMs function in the germline

PUM1 and PUM2 are expressed among most tissue types, arguably with some enrichment in the brain and heart (159, 187). Mouse PUM1 and PUM2 are expressed in fetal and adult hematopoietic stem cells and in mature blood cells (202). It has also been shown that human PUM2 is expressed in embryonic stem cells and germ cells, which contributes to the argument that the PUF proteins' primordial function lies in maintaining the germline (203).

PUFs in model organisms have long been known to be involved in germline stem cell maintenance (147-149, 151, 204, 205). FBF in *C. elegans* promotes commitment to the mitotic cell cycle by repressing the gld-1 mRNA, which promotes meiosis, thereby controlling proliferation of germline stem cells (151). Furthermore, *C. elegans* FBF represses the fem-3 mRNA, ultimately regulating the sperm-oocyte differentiation switch (166).

A gene trap approach was used to inactivate mouse PUM2 by introducing insertional mutations to produce a non-functional protein (162). This resulted in reduced testes size, but did not indicate that PUM2 is essential for reproduction or viability: the authors thusly speculated that PUM2 is required for normal testicular development (162). In a similar vein, coimmunoprecipitation experiments showed that the proteins DAZ, which is required for spermatogenesis in males, and DAZL are expressed in germ cells bind to human PUM2 (203, 206, 207). Furthermore, the protein BOL, which in fly is required for progression through meiosis in males, forms a complex with PUM2 (208). Such complexes appear to be RNA dependent, indicating that the proteins co-occupy the same RNAs (208). More evidence that PUM1 plays a role in spermatogenesis was presented recently in a study in which PUM1 null males showed a reduction in sperm count and fertility and regulate components in the p53 pathway (161).

Together, these studies highlight the likelihood that PUMs play a role in the germline, but functional evidence and studies inclusive of regulation by both PUM proteins are needed. Overlapping function of PUM1 and PUM2 may complicate the interpretation of these experiments: knockout of *both* proteins is required to fully assess their functions in mammals. In Chapters 3 and 4, I present data showing that both PUM1 and PUM2 are repressors that bind the

same sequence, further strengthening the argument that both PUMs must be considered in future experimental inquiries.

PUMs are involved in the nervous system

The central nervous system (CNS) is composed of intricate networks of different cell types transmitting information to one another in a complex and well-organized manner via synapses. Specialized cells in the brain receive and respond to signals from the environment and from within the body. Healthy cognitive function is dependent on highly regulated signaling between cell types to form dynamic networks of information transfer.

Given the complexity and dynamic nature of neural signaling, it becomes quite clear that strictly timed production of genes and their regulated decay is essential for healthy brain and CNS function. RNA binding factors play key roles in regulation of genes expressed in the nervous system. Thusly, a large number of neurological diseases are caused by defects in RBP function, mRNA expression, and translation. For example, Fragile X syndrome is caused by a loss of function of FMR1, a gene encoding the RBP FMRP: this mutation is a CGG tri-nucleotide repeat expansion and can be repeated more than 200 times in the promoter region of the gene (209, 210). Reduction of FMRP in the brain results in a number of neurobiological changes that are thought to be due to changes in synaptic plasticity and structure, including defects in learning and memory, autism-like symptoms, and hyperactivity (209). This is not a unique case, as many RBPs function in the nervous system to control mRNA localization and translation in synaptic function (211). The PUF family of proteins has been implicated in neurological function.

A large body of work implicates *Drosophila* Pumilio in neural function, observations that laid the foundation for examining mammalian PUMs in the nervous system. A number of

Drosophila protein mutants including Pumilio that disrupt long term memory but have normal short term memory were identified in a genetic screen and DNA microarray analysis (152). *In vivo* experiments indicate that Pum mutants cause neuronal excitability by way of negative regulation of para, which encodes a voltage gated Na+ channel and conversely, that para overexpression results in loss of function of Pum which can be rescued with a full length Pumilio transgene (154). Genetic experiments in female flies indicate that Pum partially rescues the sterility and other defects in neuronal function caused by a mutation in the bem gene (158). In *Drosophila* 3rd instar larvae, Pum affects the neuromuscular junction by binding the eIF4E 3' UTR to reduce its accumulation. Further, Pum localizes to the postsynaptic side of the synapse and in loss of function mutants, synaptic boutons are larger and fewer in number: together this data implicates Pum in pre- and post-synaptic function (155).

Mammalian PUM2 functions in the nervous system (159, 201). Fluorescence microscopy identified that PUM2 is found as part of mRNP complexes in rat hippocampal neurons and, under conditions of stress by arsenic treatment, localizes to dendritic stress granules (159). In later studies in rats, it was shown that loss of PUM2 in developing neurons affects dendritic outgrowth, arborization, and synaptic function (201). In the same study eIF4E mRNA was identified by immunoprecipitation with PUM2 and their interaction confirmed by gel gift assays (201). Further, eIF4E protein levels were increased by PUM2 knockdown, which provides some evidence that eIF4E is repressed by PUM2 (201).

PUM2 deficient mice exhibit abnormal behavior, increased locomotor activity, decreased body weight, low seizure thresholds, and significant changes to neuronal ion channel expression; all of which provide substantial evidence for a role for PUM2 in the nervous system (212). More recent work found that PUM2 is localized to the mouse neuromuscular junction and that PUM2

knockdown by shRNA transfection led to upregulation of the neurotransmitter acetylcholinesterase (AChE) (213). Recent studies show that PUM2 binds the voltage gated Na^+ channel transcript ($Na_V1.6$) to regulate Na^+ ion channels in neurons: loss of PUM2 results in increased current across Na channels and increased numbers of action potentials (214).

One intriguing example of PUM function in the nervous system is a recent study looking at traumatic brain injury in mice. Upon brain injury, astrocytes (a type of glial cell) undergo a process called reactive astrogliosis, which is thought to occur in response to increased Ca^{2+} levels and has both positive and negative affects on surrounding injured tissues. During this process, N-cadherin, which, owing to its 3'UTR PRE, the authors speculate is normally translationally repressed by PUM2, is upregulated in response to high Ca^{2+} and downregulated PUM2 expression (215). They demonstrate PUM regulation of N-cadherin using a reporter system, and depletion of the PUM2 RBD abrogates regulation of N-cadherin in their study (215). The importance of PUM regulation in the nervous system is evident and future work should focus on providing functional evidence of individual PUM repression of specific targets in neurons.

PUMs, miRNAs, and cell proliferation

Recent work suggests that PUMs interact functionally with the miRNA regulatory pathway (163, 170, 188, 194). p27 downregulation is required for entry into the cell cycle from quiescent cells and the authors suggest that a structural change in the p27 mRNA induced by PUM1 binding, which renders miR221/222 sites accessible for binding and repression, was responsible for miRNA mediated repression (163). Their data does point to a role for human PUM1 in cell proliferation, which is consistent with PUM function in other studies and PUF function in model organisms (161, 194). In the same vein, the E2F3 oncogene, which is typically

overexpressed in cancers, is reported to be a target of concerted repressive action between miR-503 and human PUMs in bladder cells (194). Thus, in bladder carcinomas in which PUMs and miRNAs are downregulated, repression of E2F3 is significantly reduced and the protein is upregulated. The authors postulate a similar mechanisms used by PUMs to open up secondary structure to render the miRNA site accessible by miR-503 and vulnerable to miRNA-mediated repression. Given the sheer number of miRNA sites already identified in mRNAs, PUMs and miRNAs likely share many targets. PUFs and miRNAs share corepressors: both recruit members of the CCR4-NOT deadenylase complex to enhance deadenylation and subsequent decay of mRNA targets. Thus, miRNA and PUM pathways in humans may cooperate to confer regulation of mRNA targets.

Discovery of the mechanism of repression by PUMs is essential to understand their biological functions

The work outlined in this dissertation examines the mechanisms by which human PUMs repress their target RNAs. Little is known about which human biological processes are regulated by PUMs, and even less is known about how they go about doing so. In this dissertation, I completed work with the ultimate goal of understanding the RNA targets and mechanisms of human PUM repression.

I will describe tools to address mechanistic questions related to PUM repression in Chapter 2. In Chapter 3, I will demonstrate use of the tools developed in the previous chapter to address the hypothesis that PUMs are repressors and to study mechanisms of human PUM repression. Specifically, I will show that PUM1 and PUM2 repress a luciferase reporter containing 3 PREs in its 3' UTR. In Chapter 4, I will discuss the identification of high confidence targets of human PUMs using a deep sequencing technique, RNA-Seq, and will

compare the results from that study with those already published. It is my hope that the results of this study illuminate the biological significance of PUM regulation, and provide insights into some of the mechanisms by which they serve as regulators of gene regulation.

References

1. Crick, F. H., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961) General nature of the genetic code for proteins. *Nature* **192**, 1227–1232
2. Saguez, C., Olesen, J. R., and Jensen, T. H. (2005) Formation of export-competent mRNP: escaping nuclear destruction. *Current Opinion in Cell Biology* **17**, 287–293
3. Garneau, N. L., Wilusz, J., and Wilusz, C. J. (2007) The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8**, 113–126
4. Decker, K. B., and Hinton, D. M. (2013) Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annu. Rev. Microbiol.* **67**, 113–139
5. Elkon, R., Ugalde, A. P., and Agami, R. (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nature Reviews Genetics* **14**, 496–506
6. Wu, X., and Brewer, G. (2012) The regulation of mRNA stability in mammalian cells: 2.0. *Gene* **500**, 10–21
7. Kapp, L. D., and Lorsch, J. R. (2004) The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem.* **73**, 657–704
8. Liu, X., Bushnell, D. A., and Kornberg, R. D. (2013) RNA polymerase II transcription: structure and mechanism. *Biochim. Biophys. Acta* **1829**, 2–8
9. Kuehner, J. N., Pearson, E. L., and Moore, C. (2011) Unravelling the means to an end: RNA polymerase II transcription termination. *Nat Rev Mol Cell Biol*, 1–12
10. Coppola, J. A., Field, A. S., and Luse, D. S. (2010) Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc Natl Acad Sci USA* **80**, 1251–1255
11. Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S., and Darnell, J. E. (1980) The addition of 5' cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped. *Cell* **19**, 69–78
12. Green, M. R., Maniatis, T., and Melton, D. A. (1983) Human beta-globin pre-mRNA synthesized in vitro is accurately spliced in Xenopus oocyte nuclei. *Cell* **32**, 681–694
13. Edery, I., and Sonenberg, N. (1985) Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc Natl Acad Sci USA* **82**, 7590–7594
14. Ohno, M., Sakamoto, H., and Shimura, Y. (1987) Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure. *Proc Natl Acad Sci USA* **84**, 5187–5191
15. Beelman, C. A., and Parker, R. (1994) Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA. *J. Biol. Chem.* **269**, 9687–9692
16. Jarmolowski, A., Boelens, W. C., Izaurralde, E., and Mattaj, I. W. (1994) Nuclear export of different classes of RNA is mediated by specific factors. *The Journal of Cell Biology* **124**, 627–635
17. Muthukrishnan, S., Both, G. W., Furuichi, Y., and Shatkin, A. J. (1975) 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* **255**, 33–37
18. Adams, J. M., and Cory, S. (1975) Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA. *Nature* **255**, 28–33
19. Furuichi, Y., Morgan, M., Shatkin, A. J., Jelinek, W., Salditt-Georgieff, M., and Darnell, J. E. (1975) Methylated, blocked 5 termini in HeLa cell mRNA. *Proc Natl Acad Sci USA*

72, 1904–1908
20. Furuichi, Y., LaFiandra, A., and Shatkin, A. J. (1977) 5'-Terminal structure and mRNA stability. *Nature* **266**, 235–239
21. Sonenberg, N., Morgan, M. A., Merrick, W. C., and Shatkin, A. J. (1978) A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci USA* **75**, 4843–4847
22. Topisirovic, I., Svitkin, Y. V., Sonenberg, N., and Shatkin, A. J. (2010) Cap and cap-binding proteins in the control of gene expression. *WIREs RNA* **2**, 277–298
23. Millevoi, S., and Vagner, S. (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Research* **38**, 2757–2774
24. Mandel, C. R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J. L., and Tong, L. (2006) Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**, 953–956
25. Shi, Y., Chan, S., and Martinez-Santibañez, G. (2009) An up-close look at the pre-mRNA 3'-end processing complex. *RNA Biol* **6**, 522–525
26. Shatkin, A. J., and Manley, J. L. (2000) The ends of the affair: capping and polyadenylation. *Nat. Struct. Biol.* **7**, 838–842
27. Proudfoot, N. J., and Longley, J. I. (1976) The 3' terminal sequences of human alpha and beta globin messenger RNAs: comparison with rabbit globin messenger RNA. *Cell* **9**, 733–746
28. Lutz, C. S. (2008) Alternative polyadenylation: a twist on mRNA 3' end formation. *ACS Chem. Biol.* **3**, 609–617
29. Logan, J., Falck-Pedersen, E., Darnell, J. E., and Shenk, T. (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc Natl Acad Sci USA* **84**, 8306–8310
30. Whitelaw, E., and Proudfoot, N. (1986) Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. *EMBO J* **5**, 2915–2922
31. Edwalds-Gilbert, G., Prescott, J., and Falck-Pedersen, E. (1993) 3' RNA processing efficiency plays a primary role in generating termination-competent RNA polymerase II elongation complexes. *Molecular and Cellular Biology* **13**, 3472–3480
32. Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A., and Burge, C. B. (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**, 1643–1647
33. Tian, B., Hu, J., Zhang, H., and Lutz, C. S. (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Research* **33**, 201–212
34. Wilusz, J. E., and Spector, D. L. (2010) An unexpected ending: Noncanonical 3' end processing mechanisms. *RNA* **16**, 259–266
35. Kühn, U., and Wahle, E. (2004) Structure and function of poly(A) binding proteins. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1678**, 67–84
36. Kuhn, U., Gundel, M., Knoth, A., Kerwitz, Y., Rudel, S., and Wahle, E. (2009) Poly(A) Tail Length Is Controlled by the Nuclear Poly(A)-binding Protein Regulating the Interaction between Poly(A) Polymerase and the Cleavage and Polyadenylation Specificity Factor. *Journal of Biological Chemistry* **284**, 22803–22814
37. Brook, M., and Gray, N. K. (2012) The role of mammalian poly(A)-binding proteins in

co-ordinating mRNA turnover. *Biochem. Soc. Trans* **40**, 856–864

38. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., la Bastide, de, M., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y. J., Szustakowki, J., International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921

39. Schröder, H. C., Trölltsch, D., Friese, U., Bachmann, M., and Müller, W. E. (1987) Mature mRNA is selectively released from the nuclear matrix by an ATP/dATP-dependent mechanism sensitive to topoisomerase inhibitors. *J. Biol. Chem.* **262**, 8917–

8925
40. Eckner, R., Ellmeier, W., and Birnstiel, M. L. (1991) Mature mRNA 3' end formation stimulates RNA export from the nucleus. *EMBO J* **10**, 3513–3522
41. Goldstrohm, A. C., and Wickens, M. (2008) Multifunctional deadenylase complexes diversify mRNA control. *Nat Rev Mol Cell Biol* **9**, 337–344
42. Preiss, T., and Hentze, M. W. (1998) Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. *Nature* **392**, 516–520
43. Gallie, D. R. (1991) The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes & Development* **5**, 2108–2116
44. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* **473**, 337–342
45. Aitken, C. E., and Lorsch, J. R. (2012) A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol* **19**, 568–576
46. Kozak, M. (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *Journal of Molecular Biology* **196**, 947–950
47. Kozak, M. (1980) Binding of wheat germ ribosomes to bisulfite-modified reovirus messenger RNA: evidence for a scanning mechanism. *Journal of Molecular Biology* **144**, 291–304
48. Kozak, M. (1980) Role of ATP in binding and migration of 40S ribosomal subunits. *Cell* **22**, 459–467
49. Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research* **15**, 8125–8148
50. Wickens, N. K. G. A. M. (1998) CONTROL OF TRANSLATION INITIATION IN ANIMALS. 1–60
51. Kozak, M. (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Molecular and Cellular Biology* **9**, 5134–5142
52. Richter, J. D. (2007) CPEB: a life in translation. *Trends in Biochemical Sciences* **32**, 279–285
53. Kong, J., and Lasko, P. (2012) Translational control in cellular and developmental processes. *Nature Reviews Genetics* **13**, 383–394
54. Miller, M. A., and Olivas, W. M. (2011) Roles of Puf proteins in mRNA degradation and translation. *WIREs RNA* **2**, 471–492
55. Yang, E., van Nimwegen, E., Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., and Darnell, J. E. (2003) Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. *Genome Research* **13**, 1863–1872
56. Sharova, L. V., Sharov, A. A., Nedorezov, T., Piao, Y., Shaik, N., and Ko, M. S. H. (2009) Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res.* **16**, 45–58
57. Raghavan, A., Ogilvie, R. L., Reilly, C., Abelson, M. L., Raghavan, S., Vasdewani, J., Krathwohl, M., and Bohjanen, P. R. (2002) Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes. *Nucleic Acids Research* **30**, 5529–5538
58. Shyu, A.-B., Belasco, J. G., and Greenberg, M. E. (1991) Two distinct destabilizing

elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. *Genes & Development* **5**, 221–231
59. Muhlrad, D., and Parker, R. (1992) Mutations affecting stability and deadenylation of the yeast MFA2 transcript. *Genes & Development* **6**, 2100–2111
60. Charlesworth, A., Meijer, H. A., and de Moor, C. H. (2013) Specificity factors in cytoplasmic polyadenylation. *WIREs RNA* **4**, 437–461
61. Zuo, Y., and Deutscher, M. P. (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Research* **29**, 1017–1026
62. Dlakić, M. (2000) Functionally unrelated signalling proteins contain a fold similar to Mg2+-dependent endonucleases. *Trends in Biochemical Sciences* **25**, 272–273
63. Yamashita, A., Chang, T.-C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.-Y. A., and Shyu, A.-B. (2005) Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol* **12**, 1054–1063
64. Brown, C. E., Tarun, S. Z., Boeck, R., and Sachs, A. B. (1996) PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **16**, 5744–5753
65. Tucker, M., Valencia-Sanchez, M. A., Staples, R. R., Chen, J., Denis, C. L., and Parker, R. (2001) The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. *Cell* **104**, 377–386
66. Tucker, M., Staples, R. R., Valencia-Sanchez, M. A., Muhlrad, D., and Parker, R. (2002) Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. *EMBO J* **21**, 1427–1436
67. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3“-->5” exoribonucleases. *Cell* **91**, 457–466
68. Anderson, J. S., and Parker, R. P. (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3“ to 5” exonucleases of the exosome complex. *EMBO J* **17**, 1497–1506
69. Denis, C. L., and Chen, J. (2003) The CCR4-NOT complex plays diverse roles in mRNA metabolism. *Prog. Nucleic Acid Res. Mol. Biol.* **73**, 221–250
70. Morel, A.-P., Sentis, S., Bianchin, C., Le Romancer, M., Jonard, L., Rostan, M.-C., Rimokh, R., and Corbo, L. (2003) BTG2 antiproliferative protein interacts with the human CCR4 complex existing in vivo in three cell-cycle-regulated forms. *Journal of Cell Science* **116**, 2929–2936
71. Doidge, R., Mittal, S., Aslam, A., and Winkler, G. S. (2012) Deadenylation of cytoplasmic mRNA by the mammalian Ccr4-Not complex. *Biochem. Soc. Trans* **40**, 896–901
72. Collart, M. A., and Panasenko, O. O. (2012) The Ccr4--not complex. *Gene* **492**, 42–53
73. Draper, M. P., Salvadore, C., and Denis, C. L. (1995) Identification of a mouse protein whose homolog in Saccharomyces cerevisiae is a component of the CCR4 transcriptional regulatory complex. *Molecular and Cellular Biology* **15**, 3487–3495
74. Viswanathan, P., Ohn, T., Chiang, Y.-C., Chen, J., and Denis, C. L. (2004) Mouse CAF1 can function as a processive deadenylase/3“-5-”exonuclease in vitro but in yeast the deadenylase function of CAF1 is not required for mRNA poly(A) removal. *J. Biol. Chem.* **279**, 23988–23995

75. Shimizu-Yoshida, Y., Sasamoto, M., Yoshida, A., Yoshioka, T., Matsumoto, A., and Sakai, A. (1999) Mouse CAF1, a mouse homologue of the yeast POP2 gene, complements the yeast pop2 null mutation. *Yeast* **15**, 1357–1364
76. Schwede, A., Ellis, L., Luther, J., Carrington, M., Stoecklin, G., and Clayton, C. (2008) A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Research* **36**, 3374–3388
77. Albert, T. K., Lemaire, M., van Berkum, N. L., Gentz, R., Collart, M. A., and Timmers, H. T. (2000) Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. *Nucleic Acids Research* **28**, 809–817
78. Muhlrad, D., Decker, C. J., and Parker, R. (1994) Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5“-->3” digestion of the transcript. *Genes & Development* **8**, 855–866
79. Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007) The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* **13**, 998–1016
80. Tharun, S. (2009) Lsm1-7-Pat1 complex: a link between 3“ and 5-”ends in mRNA decay? *RNA Biol* **6**, 228–232
81. Song, M.-G., Li, Y., and Kiledjian, M. (2010) Multiple mRNA Decapping Enzymes in Mammalian Cells. *Molecular Cell* **40**, 423–432
82. Schwartz, D. C., and Parker, R. (2000) mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Molecular and Cellular Biology* **20**, 7933–7942
83. Arribas-Layton, M., Wu, D., Lykke-Andersen, J., and Song, H. (2013) Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochim. Biophys. Acta* **1829**, 580–589
84. Amrani, N., Ghosh, S., Mangus, D. A., and Jacobson, A. (2008) Translation factors promote the formation of two states of the closed-loop mRNP. *Nature* **453**, 1276–1280
85. Wilson, T., and Treisman, R. (1988) Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3' AU-rich sequences. *Nature* **336**, 396–399
86. Decker, C. J., and Parker, R. (1993) A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes & Development* **7**, 1632–1643
87. Muhlrad, D., Decker, C. J., and Parker, R. (1995) Turnover mechanisms of the stable yeast PGK1 mRNA. *Molecular and Cellular Biology* **15**, 2145–2156
88. Gallie, D. R., and Tanguay, R. (1994) Poly(A) binds to initiation factors and increases cap-dependent translation in vitro. *J. Biol. Chem.* **269**, 17166–17173
89. Schwartz, D. C., and Parker, R. (1999) Mutations in translation initiation factors lead to increased rates of deadenylation and decapping of mRNAs in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **19**, 5247–5256
90. Baltz, A. G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., and Landthaler, M. (2012) The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Molecular Cell* **46**, 674–690
91. Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., Davey, N. E., Humphreys, D. T., Preiss, T., Steinmetz, L. M., Krijgsveld, J., and Hentze,

M. W. (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, 1–24

92. Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L. H., Dale, R. K., Smith, S. A., Yarosh, C. A., Kelly, S. M., Nabet, B., Mecenas, D., Li, W., Laishram, R. S., Qiao, M., Lipshitz, H. D., Piano, F., Corbett, A. H., Carstens, R. P., Frey, B. J., Anderson, R. A., Lynch, K. W., Penalva, L. O. F., Lei, E. P., Fraser, A. G., Blencowe, B. J., Morris, Q. D., and Hughes, T. R. (2014) A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177
93. Mitchell, S. F., Jain, S., She, M., and Parker, R. (2012) Global analysis of yeast mRNPs. *Nat Struct Mol Biol*, 1–9
94. Castello, A., Fischer, B., Hentze, M. W., and Preiss, T. (2013) RNA-binding proteins in Mendelian disease. *Trends Genet.* **29**, 318–327
95. Frenquelli, M., Muzio, M., Scielzo, C., Fazi, C., Scarfo, L., Rossi, C., Ferrari, G., Ghia, P., and Caligaris-Cappio, F. (2010) MicroRNA and proliferation control in chronic lymphocytic leukemia: functional relationship between miR-221/222 cluster and p27. *Blood* **115**, 3949–3959
96. Esquela-Kerscher, A., and Slack, F. J. (2006) Oncomirs — microRNAs with a role in cancer. *Nat Rev Cancer* **6**, 259–269
97. Osborne, R. J. (2006) RNA-dominant diseases. *Human Molecular Genetics* **15**, R162–R169
98. Kraut-Cohen, J., and Gerst, J. E. (2010) Addressing mRNAs to the ER: cis sequences act up! *Trends in Biochemical Sciences* **35**, 459–469
99. LaGrandeur, T., and Parker, R. (1999) The cis acting sequences responsible for the differential decay of the unstable MFA2 and stable PGK1 transcripts in yeast include the context of the translational start codon. *RNA* **5**, 420–433
100. Listowski, M. A., Heger, E., Bogusławska, D. M., Machnicka, B., Kuliczkowski, K., Leluk, J., and Sikorski, A. F. (2013) microRNAs: fine tuning of erythropoiesis. *Cell. Mol. Biol. Lett.* **18**, 34–46
101. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854
102. Wightman, B., Ha, I., and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75**, 855–862
103. Bartel, D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297
104. Lee, Y., Jeon, K., Lee, J.-T., Kim, S., and Kim, V. N. (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* **21**, 4663–4670
105. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, V. N. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419
106. Yi, R., Qin, Y., Macara, I. G., and Cullen, B. R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development* **17**, 3011–3016
107. Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. (2004) Nuclear export of microRNA precursors. *Science* **303**, 95–98

108. Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* **106**, 23–34
109. Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, E., Tuschl, T., and Zamore, P. D. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834–838
110. Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes & Development* **15**, 2654–2659
111. Wilson, R. C., and Doudna, J. A. (2013) Molecular Mechanisms of RNA Interference. *Annu. Rev. Biophys.* **42**, 217–239
112. Pasquinelli, A. E. (2012) MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nature Reviews Genetics* **13**, 271–282
113. Loeb, G. B., Khan, A. A., Canner, D., Hiatt, J. B., Shendure, J., Darnell, R. B., Leslie, C. S., and Rudensky, A. Y. (2012) Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. *Molecular Cell* **48**, 760–770
114. Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* **122**, 553–563
115. Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005) Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **309**, 1573–1576
116. Wang, B., Yanez, A., and Novina, C. D. (2008) MicroRNA-repressed mRNAs contain 40S but not 60S components. *Proceedings of the National Academy of Sciences* **105**, 5343–5348
117. Humphreys, D. T., Westman, B. J., Martin, D. I. K., and Preiss, T. (2005) MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci USA* **102**, 16961–16966
118. Mathonnet, G., Fabian, M. R., Svitkin, Y. V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W. C., Darzynkiewicz, E., Pillai, R. S., Filipowicz, W., Duchaine, T. F., and Sonenberg, N. (2007) MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science* **317**, 1764–1767
119. Thermann, R., and Hentze, M. W. (2007) Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **447**, 875–878
120. Olsen, P. H., and Ambros, V. (1999) The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* **216**, 671–680
121. Nottrott, S., Simard, M. J., and Richter, J. D. (2006) Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat Struct Mol Biol* **13**, 1108–1114
122. Petersen, C. P., Bordeleau, M.-E., Pelletier, J., and Sharp, P. A. (2006) Short RNAs repress translation after initiation in mammalian cells. *Molecular Cell* **21**, 533–542
123. Maroney, P. A., Yu, Y., Fisher, J., and Nilsen, T. W. (2006) Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat Struct Mol Biol* **13**, 1102–1107

124. Wu, L., Fan, J., and Belasco, J. G. (2006) MicroRNAs direct rapid deadenylation of mRNA. *Proc Natl Acad Sci USA* **103**, 4034–4039
125. Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006) mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & Development* **20**, 1885–1898
126. Chen, C.-Y. A., Zheng, D., Xia, Z., and Shyu, A.-B. (2009) Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nat Struct Mol Biol* **16**, 1160–1166
127. Huntzinger, E., Braun, J. E., Heimstädt, S., Zekri, L., and Izaurralde, E. (2010) Two PABPC1-binding sites in GW182 proteins promote miRNA-mediated gene silencing. *EMBO J* **29**, 4146–4160
128. Fabian, M. R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J. T., Svitkin, Y. V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J. A., Chen, C.-Y. A., Shyu, A.-B., Yates, J. R., Hannon, G. J., Filipowicz, W., Duchaine, T. F., and Sonenberg, N. (2009) Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Molecular Cell* **35**, 868–880
129. Zekri, L., rk, D. K. G. L.-&. Z. U., and Izaurralde, E. (2013) GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J* **32**, 1052–1065
130. Shaw, G., and Kamen, R. (1986) A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46**, 659–667
131. Schott, J., and Stoecklin, G. (2010) Networks controlling mRNA decay in the immune system. *WIREs RNA* **1**, 432–456
132. Barreau, C., Paillard, L., and Osborne, H. B. (2005) AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Research* **33**, 7138–7150
133. Chen, C. Y., Gherzi, R., Ong, S. E., Chan, E. L., Raijmakers, R., Pruijn, G. J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* **107**, 451–464
134. Fenger-Grøn, M., Fillman, C., Norrild, B., and Lykke-Andersen, J. (2005) Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Molecular Cell* **20**, 905–915
135. Lykke-Andersen, J., and Wagner, E. (2005) Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes & Development* **19**, 351–361
136. Fabian, M. R., Frank, F., Rouya, C., Siddiqui, N., Lai, W. S., Karetnikov, A., Blackshear, P. J., Nagar, B., and Sonenberg, N. (2013) Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol*
137. Ogilvie, R. L., Abelson, M., Hau, H. H., Vlasova, I., Blackshear, P. J., and Bohjanen, P. R. (2005) Tristetraprolin down-regulates IL-2 gene expression through AU-rich element-mediated mRNA decay. *J. Immunol.* **174**, 953–961
138. Levy, N. S., Chung, S., Furneaux, H., and Levy, A. P. (1998) Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR. *J. Biol. Chem.* **273**, 6417–6423
139. Mukherjee, N., Corcoran, D. L., Nusbaum, J. D., Reid, D. W., Georgiev, S., Hafner, M.,

Ascano, M., Tuschl, T., Ohler, U., and Keene, J. D. (2011) Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Molecular Cell* **43**, 327–339

140. Lebedeva, S., Jens, M., Theil, K., Schwanhausser, B., Selbach, M., Landthaler, M., and Rajewsky, N. (2011) Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Molecular Cell* **43**, 340–352
141. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141
142. Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997) The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**, 1421–1433
143. Wharton, R. P., and Struhl, G. (1991) RNA regulatory elements mediate control of Drosophila body pattern by the posterior morphogen nanos. *Cell* **67**, 955–967
144. Murata, Y., and Wharton, R. P. (1995) Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in Drosophila embryos. *Cell* **80**, 747–756
145. Wharton, R. P., Sonoda, J., Lee, T., Patterson, M., and Murata, Y. (1998) The Pumilio RNA-binding domain is also a translational regulator. *Molecular Cell* **1**, 863–872
146. Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickens, M. P. (1997) A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* **390**, 477–484
147. Lin, H., and Spradling, A. C. (1997) A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development* **124**, 2463–2476
148. Parisi, M., and Lin, H. (1999) The Drosophila pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis. *Genetics* **153**, 235–250
149. Asaoka-Taguchi, M., Yamada, M., Nakamura, A., Hanyu, K., and Kobayashi, S. (1999) Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos. *Nature Publishing Group* **1**, 431–437
150. Barker, D. D., Wang, C., Moore, J., Dickinson, L. K., and Lehmann, R. (1992) Pumilio is essential for function but not for distribution of the Drosophila abdominal determinant Nanos. *Genes & Development* **6**, 2312–2326 [online] http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1459455&retmode=ref&cmd=prlinks.
151. Crittenden, S. L., Bernstein, D. S., Bachorik, J. L., Thompson, B. E., Gallegos, M., Petcherski, A. G., Moulder, G., Barstead, R., Wickens, M., and Kimble, J. (2002) A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans. *Nature* **417**, 660–663
152. Dubnau, J., Chiang, A.-S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., Broger, C., and Tully, T. (2003) The staufen/pumilio pathway is involved in Drosophila long-term memory. *Current Biology* **13**, 286–296
153. Lehmann, R., and Nüsslein-Volhard, C. (1987) Involvement of the pumilio gene in the transport of an abdominal signal in the Drosophila embryo.
154. Mee, C. J., Pym, E. C. G., Moffat, K. G., and Baines, R. A. (2004) Regulation of

neuronal excitability through pumilio-dependent control of a sodium channel gene. *Journal of Neuroscience* **24**, 8695–8703

155. Menon, K. P., Sanyal, S., Habara, Y., Sanchez, R., Wharton, R. P., Ramaswami, M., and Zinn, K. (2004) The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E. *Neuron* **44**, 663–676
156. Nakahata, S., Kotani, T., Mita, K., Kawasaki, T., Katsu, Y., Nagahama, Y., and Yamashita, M. (2003) Involvement of Xenopus Pumilio in the translational regulation that is specific to cyclin B1 mRNA during oocyte maturation. *Mech. Dev.* **120**, 865–880
157. Rouhana, L., Wang, L., Buter, N., Kwak, J. E., Schiltz, C. A., Gonzalez, T., Kelley, A. E., Landry, C. F., and Wickens, M. (2005) Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA* **11**, 1117–1130
158. Schweers, B. A., Walters, K. J., and Stern, M. (2002) The Drosophila melanogaster translational repressor pumilio regulates neuronal excitability. *Genetics* **161**, 1177–1185
159. Vessey, J. P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M. A., and Macchi, P. (2006) Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. *Journal of Neuroscience* **26**, 6496–6508
160. Wreden, C., Verrotti, A. C., Schisa, J. A., Lieberfarb, M. E., and Strickland, S. (1997) Nanos and pumilio establish embryonic polarity in Drosophila by promoting posterior deadenylation of hunchback mRNA. *Development* **124**, 3015–3023
161. Chen, D., Zheng, W., Lin, A., Uyhazi, K., Zhao, H., and Lin, H. (2012) Pumilio 1 Suppresses Multiple Activators of p53 to Safeguard Spermatogenesis. *Curr Biol*
162. Xu, E. Y., Chang, R., Salmon, N. A., and Reijo Pera, R. A. (2007) 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
163. Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A. F., Elkon, R., and Agami, R. (2010) A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat Cell Biol* **12**, 1014–1020
164. Wickens, M., Bernstein, D. S., Kimble, J., and Parker, R. (2002) A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet.* **18**, 150–157
165. Zamore, P. D., Bartel, D. P., Lehmann, R., and Williamson, J. R. (1999) The PUMILIO−RNA Interaction: A Single RNA-Binding Domain Monomer Recognizes a Bipartite Target Sequence †. *Biochemistry* **38**, 596–604
166. Ahringer, J., and Kimble, J. (1991) Control of the sperm-oocyte switch in Caenorhabditis elegans hermaphrodites by the fem-3 3' untranslated region. *Nature* **349**, 346–348
167. Chagnovich, D., and Lehmann, R. (2001) Poly(A)-independent regulation of maternal hunchback translation in the Drosophila embryo. *Proc Natl Acad Sci USA* **98**, 11359–11364
168. Gamberi, C., Peterson, D. S., He, L., and Gottlieb, E. (2002) An anterior function for the Drosophila posterior determinant Pumilio. *Development* **129**, 2699–2710 [online] http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=12015297&retmode=ref&cmd=prlinks.
169. Chritton, J. J., and Wickens, M. (2010) Translational repression by PUF proteins in vitro. *RNA* **16**, 1217–1225
170. Friend, K., Campbell, Z. T., Cooke, A., Kroll-Conner, P., Wickens, M. P., and Kimble,

J. (2012) A conserved PUF-Ago-eEF1A complex attenuates translation elongation. *Nat Struct Mol Biol* **19**, 176–183
171. Cao, Q., Padmanabhan, K., and Richter, J. D. (2010) Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition. *RNA* **16**, 221–227
172. doi:10.1016/j.cub.2006.08.093 (2006) doi:10.1016/j.cub.2006.08.093. 1–7
173. Goldstrohm, A. C., Hook, B. A., Seay, D. J., and Wickens, M. (2006) PUF proteins bind Pop2p to regulate messenger RNAs. *Nat Struct Mol Biol* **13**, 533–539
174. Blewett, N. H., and Goldstrohm, A. C. (2012) A eukaryotic translation initiation factor 4E-binding protein promotes mRNA decapping and is required for PUF repression. *Molecular and Cellular Biology* **32**, 4181–4194
175. Eliyahu, E., Pnueli, L., Melamed, D., Scherrer, T., Gerber, A. P., Pines, O., Rapaport, D., and Arava, Y. (2010) Tom20 mediates localization of mRNAs to mitochondria in a translation-dependent manner. *Molecular and Cellular Biology* **30**, 284–294
176. Saint-Georges, Y., Garcia, M., Delaveau, T., Jourdren, L., Le Crom, S., Lemoine, S., Tanty, V., Devaux, F., and Jacq, C. (2008) Yeast Mitochondrial Biogenesis: A Role for the PUF RNA-Binding Protein Puf3p in mRNA Localization. *PLoS ONE* **3**, e2293
177. Garcia-Rodriguez, L. J., Gay, A. C., and Pon, L. A. (2007) Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast. *The Journal of Cell Biology* **176**, 197–207
178. Deng, Y., Singer, R. H., and Gu, W. (2008) Translation of ASH1 mRNA is repressed by Puf6p-Fun12p/eIF5B interaction and released by CK2 phosphorylation. *Genes & Development* **22**, 1037–1050
179. Suh, N., Crittenden, S. L., Goldstrohm, A., Hook, B., Thompson, B., Wickens, M., and Kimble, J. (2009) FBF and its dual control of gld-1 expression in the Caenorhabditis elegans germline. *Genetics* **181**, 1249–1260
180. Edwards, T. A., Pyle, S. E., Wharton, R. P., and Aggarwal, A. K. (2001) Structure of Pumilio reveals similarity between RNA and peptide binding motifs. *Cell* **105**, 281–289
181. Wang, X., McLachlan, J., Zamore, P. D., and Hall, T. M. T. (2002) Modular recognition of RNA by a human pumilio-homology domain. *Cell* **110**, 501–512
182. Wang, X., Zamore, P. D., and Hall, T. M. (2001) Crystal structure of a Pumilio homology domain. *Molecular Cell* **7**, 855–865
183. Filipovska, A., and Rackham, O. (2011) Designer RNA-binding proteins: New tools for manipulating the transcriptome. *RNA Biol* **8**, 978–983
184. Lu, G., and Hall, T. M. T. (2011) Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure* **19**, 361–367
185. Cheong, C.-G., and Hall, T. M. T. (2006) Engineering RNA sequence specificity of Pumilio repeats. *Proc Natl Acad Sci USA* **103**, 13635–13639
186. Filipovska, A., Razif, M. F. M., Nygård, K. K. A., and Rackham, O. (2011) A universal code for RNA recognition by PUF proteins. *Nat Chem Biol* **7**, 425–427
187. Spassov, D. S., and Jurecic, R. (2002) Cloning and comparative sequence analysis of PUM1 and PUM2 genes, human members of the Pumilio family of RNA-binding proteins. *Gene* **299**, 195–204
188. Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., and Gerber, A. P. (2008) Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System. *PLoS ONE* **3**, e3164

189. Morris, A. R., Mukherjee, N., and Keene, J. D. (2008) Ribonomic Analysis of Human Pum1 Reveals cis-trans Conservation across Species despite Evolution of Diverse mRNA Target Sets. *Molecular and Cellular Biology* **28**, 4093–4103
190. Keene, J. D., Komisarow, J. M., and Friedersdorf, M. B. (2006) RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc* **1**, 302–307
191. Gerber, A. P., Herschlag, D., and Brown, P. O. (2004) Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *Plos Biol* **2**, e79 [online] http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15024427&retmode=ref&cmd=prlinks.
192. Gerber, A. P., Luschnig, S., Krasnow, M. A., Brown, P. O., and Herschlag, D. (2006) Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in Drosophila melanogaster. *Proc Natl Acad Sci USA* **103**, 4487–4492
193. Lee, M.-H., Hook, B., Pan, G., Kershner, A. M., Merritt, C., Seydoux, G., Thomson, J. A., Wickens, M., and Kimble, J. (2007) Conserved regulation of MAP kinase expression by PUF RNA-binding proteins. *PLoS Genet* **3**, e233
194. Miles, W. O., Tschöp, K., Herr, A., Ji, J.-Y., and Dyson, N. J. (2012) Pumilio facilitates miRNA regulation of the E2F3 oncogene. *Genes & Development* **26**, 356–368
195. Goldstrohm, A. C., Seay, D. J., Hook, B. A., and Wickens, M. (2007) PUF protein-mediated deadenylation is catalyzed by Ccr4p. *J. Biol. Chem.* **282**, 109–114
196. Ahringer, J., Rosenquist, T. A., Lawson, D. N., and Kimble, J. (1992) The Caenorhabditis elegans sex determining gene fem-3 is regulated post-transcriptionally. *EMBO J* **11**, 2303–2310
197. Gamberi, C., Peterson, D. S., He, L., and Gottlieb, E. (2002) An anterior function for the Drosophila posterior determinant Pumilio. *Development* **129**, 2699–2710
198. Olivas, W., and Parker, R. (2000) The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. *EMBO J* **19**, 6602–6611
199. Hook, B. A., Goldstrohm, A. C., Seay, D. J., and Wickens, M. (2007) Two yeast PUF proteins negatively regulate a single mRNA. *J. Biol. Chem.* **282**, 15430–15438
200. Lehmann, R., and Nüsslein-Volhard, C. (1991) The maternal gene nanos has a central role in posterior pattern formation of the Drosophila embryo. *Development* **112**, 679–691
201. Vessey, J. P., Schoderboeck, L., Gingl, E., Luzi, E., Riefler, J., Di Leva, F., Karra, D., Thomas, S., Kiebler, M. A., and Macchi, P. (2010) Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function. *Proceedings of the National Academy of Sciences* **107**, 3222–3227
202. Spassov, D. S., and Jurecic, R. (2003) Mouse Pum1 and Pum2 genes, members of the Pumilio family of RNA-binding proteins, show differential expression in fetal and adult hematopoietic stem cells and progenitors. *Blood Cells Mol. Dis.* **30**, 55–69
203. Moore, F. L., Jaruzelska, J., Fox, M. S., Urano, J., Firpo, M. T., Turek, P. J., Dorfman, D. M., and Pera, R. A. R. (2003) Human Pumilio-2 is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-like proteins. *Proc Natl Acad Sci USA* **100**, 538–543
204. Forbes, A., and Lehmann, R. (1998) Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. *Development* **125**, 679–

690
205. Kraemer, B., Crittenden, S., Gallegos, M., Moulder, G., Barstead, R., Kimble, J., and Wickens, M. (1999) NANOS-3 and FBF proteins physically interact to control the sperm-oocyte switch in Caenorhabditis elegans. *Current Biology* **9**, 1009–1018
206. Reijo, R., Lee, T. Y., Salo, P., Alagappan, R., Brown, L. G., Rosenberg, M., Rozen, S., Jaffe, T., Straus, D., and Hovatta, O. (1995) Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet* **10**, 383–393
207. Reijo, R., Alagappan, R. K., Patrizio, P., and Page, D. C. (1996) Severe oligozoospermia resulting from deletions of azoospermia factor gene on Y chromosome. *Lancet* **347**, 1290–1293
208. Urano, J., Fox, M. S., and Reijo Pera, R. A. (2005) Interaction of the conserved meiotic regulators, Boule (BOL) and Pumilio-2 (PUM2). *Mol. Reprod. Dev.* **71**, 290–298
209. Pfeiffer, B. E., and Huber, K. M. (2009) The state of synapses in fragile X syndrome. *Neuroscientist* **15**, 549–567
210. Ascano, M., Mukherjee, N., Bandaru, P., Miller, J. B., Nusbaum, J. D., Corcoran, D. L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M., Williams, Z., Ohler, U., and Tuschl, T. (2012) FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* **492**, 382–386
211. Costa-Mattioli, M., Sossin, W. S., Klann, E., and Sonenberg, N. (2009) Translational Control of Long-Lasting Synaptic Plasticity and Memory. *Neuron* **61**, 10–26
212. Siemen, H., Colas, D., Heller, H. C., Brüstle, O., and Pera, R. A. R. (2011) Pumilio-2 function in the mouse nervous system. *PLoS ONE* **6**, e25932
213. Marrero, E., Rossi, S. G., Darr, A., Tsoulfas, P., and Rotundo, R. L. (2011) Translational regulation of acetylcholinesterase by the RNA-binding protein Pumilio-2 at the neuromuscular synapse. *Journal of Biological Chemistry* **286**, 36492–36499
214. Driscoll, H. E., Muraro, N. I., He, M., and Baines, R. A. (2013) Pumilio-2 regulates translation of NaV1.6 to mediate homeostasis of membrane excitability. *Journal of Neuroscience* **33**, 9644–9654
215. Kanemaru, K., Kubota, J., Sekiya, H., Hirose, K., Okubo, Y., and Iino, M. (2013) Calcium-dependent N-cadherin up-regulation mediates reactive astrogliosis and neuroprotection after brain injury. *Proceedings of the National Academy of Sciences* **110**, 11612–11617

Figures

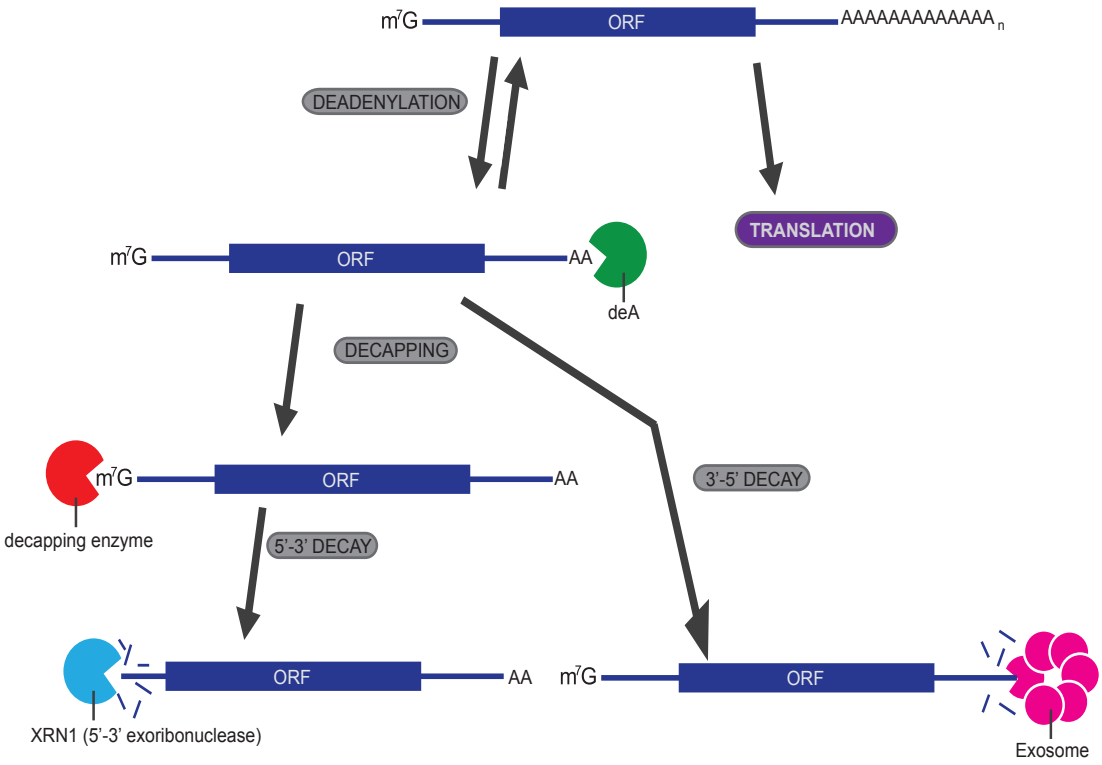
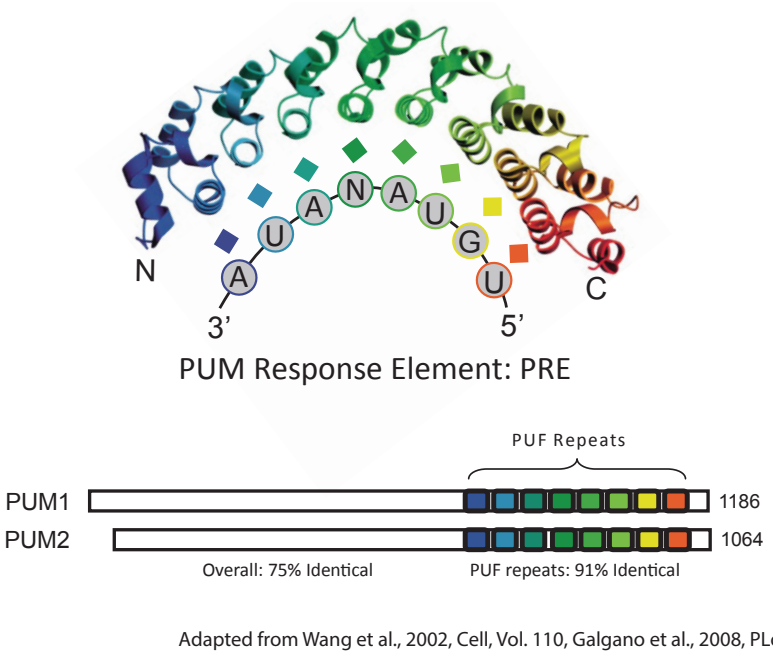


Figure 1.1 PUMs bind single stranded RNA via their RNA binding domains.

A cartoon depiction of a human PUM RNA binding domain interacting with its consensus 8 nucleotide binding sequence, termed the PUM response element. Stick diagrams of human PUM1 and PUM2 are depicted below, demonstrating the high degree of similarity between their C terminal RNA binding domains (rainbow blocks).



Adapted from Wang et al., 2002, Cell, Vol. 110, Galgano et al., 2008, PLoS ONE, Vol. 3

Figure 1.2 PUMs bind single stranded RNA via their RNA binding domains.

A cartoon depiction of a human PUM RNA binding domain interacting with its consensus 8 nucleotide binding sequence, termed the PUM response element. Stick diagrams of human PUM1 and PUM2 are depicted below, demonstrating the high degree of similarity between their C terminal RNA binding domains (rainbow blocks).

CHAPTER 2

A guide to design and optimization of reporter assays for 3' untranslated region mediated regulation of mammalian messenger RNAs

This article is currently in press: J. Van Etten et al., Methods (2013), http://dx.doi.org/10.1016/j.ymeth.2013.04.020. Trista Schagat and I contributed equally to writing the manuscript and to making figures.

Abstract

Post-transcriptional regulatory mechanisms are pervasive in the control of gene expression. Regulatory sequences within transcripts can control RNA processing, localization, translation efficiency, and stability of the RNA. Regulation is mediated by a diverse set of RNA binding regulators, including proteins and RNAs, which interact with specific mRNA sequences that are often found in untranslated regions. The potential for vast post-transcriptional control exists: mammalian mRNAs contain extensive untranslated regions and their genomes encode many hundreds of RNA binding proteins and non-coding RNAs. Facile quantitative methods are

necessary to study the activities and mechanisms of regulatory sequences and the RNA binding factors that recognize them. Here we discuss the design and implementation of luciferase-based reporter assays to measure the effect of regulatory RNA sequences on protein and RNA expression. Protocols are described for transfection of the reporter into cells, measurement of protein expression levels with luciferase activity assays, RNA purification, and measurement of mRNA levels by reverse transcription and quantitative polymerase chain reaction. For each assay, troubleshooting of common problems and critical controls are discussed. We present our optimized techniques and data from studies that measure specific and direct repression (i.e. negative regulation) of mRNAs by members of the PUF family of RNA binding proteins in cultured human cells.

Introduction

Gene expression - the complex process of decoding genetic information from DNA to RNA to protein – is controlled by numerous regulatory factors at multiple steps. Messenger RNA (mRNA) is synthesized in the nucleus where, before its export to the cytoplasm, introns are spliced out and 5' and 3' ends are capped and polyadenylated, respectively (1). Once in the cytoplasm, mRNA is translated, localized, stored, or degraded (1, 2). These cytoplasmic processes can be regulated by sequences often found in untranslated regions (UTRs) within the mRNA. Regulatory sequences control translation efficiency, RNA stability, and in some cases act as zip codes that direct the mRNA to specific intracellular locations: in many instances, trans-acting factors recognize these regulatory sequences to control the message (2, 3). Regulation can be positive or negative. Hundreds of potential regulatory proteins and small RNAs are encoded within mammalian genomes; therefore, there is enormous potential for RNA-mediated regulation

of gene expression. Quantitative methods that can measure and dissect post-transcriptional regulatory mechanisms are necessary.

Regulation of natural mRNAs can be detected using techniques such as Northern blotting or reverse transcription coupled with quantitative polymerase chain reaction (RT-qPCR). To detect post-transcriptional regulation of protein expression, Western blotting can be performed with a specific antibody, but is often limited in detection range and sensitivity. On a global scale, changes in RNA levels can be analyzed using microarrays or high-throughput RNA sequencing approaches. Likewise, global changes in protein synthesis can be assessed, for instance by labeling with amino acids containing stable isotopes coupled with relative quantitation by mass spectrometry. Yet these approaches are technically demanding and require sophisticated instrumentation. While these approaches are valuable, outside of a genetically tractable model system, dissection of the regulatory sequences in natural mRNAs can be difficult if not impractical.

Reporter genes have proven valuable to isolate regulatory elements and dissect their impact on expression of the reporter mRNA and protein. Reporters provide a quantitative readout at both the RNA and protein levels. Insertion of a regulatory sequence into the context of a reporter gene provides a powerful test for its activity. Expression can then be monitored at both the RNA and protein level to detect the elements effect on translation and mRNA stability. Reporter protein assays offer broad dynamic range of detection with increased sensitivity relative to other protein detection assays such as Western blotting. In this context, the regulatory sequence can be dissected, manipulated and mutated. In this manner, the minimal regulatory element can be delineated and mutations that alter its function can be identified. The regulatory impact of specific trans-acting factors can be investigated by depleting the factor using RNA

interference or by over-expressing a wild type or dominant negative form of the factor and then measuring its impact on reporter mRNA and protein expression. Once established, the reporter gene assay can also provide a powerful tool for screens designed to identify trans-acting regulators, responses to stimuli, and to identify molecules that can alter the regulatory response.

Here we describe a luciferase reporter gene-based approach to study mRNA regulation. This chapter is designed especially to assist researchers new to investigating post-transcriptional regulation, but the information will likely prove useful to experienced scientists. First, the regulatory sequence of interest is engineered into the appropriate location of the reporter gene. While our focus is on regulatory sequences located in the 3'UTR, the approaches are adaptable to analysis of 5'UTR elements. The reporter gene is introduced into a model cell line by transfection. To measure protein expression, luciferase activity assays are performed; the amount of light produced is directly proportional to the amount of protein expressed. To measure effects on mRNA expression, the luciferase mRNA levels are measured by RT-qPCR.

We demonstrate the utility of this approach in the study of RNA regulation by PUF proteins. PUF proteins are RNA binding proteins that recognize specific regulatory sequences typically found in the 3' UTR of certain mRNAs (3). In model organisms, they are known to repress protein expression; however, the activities of human PUF proteins, PUM1 and PUM2, were not known (3). PUM1 and PUM2 bind to the same RNA sequence, termed the Pum Response Element (PRE) (4-7). To study their function, we introduced PREs into the 3'UTR of a *Renilla* luciferase reporter gene. We found that both PUMs repress reporter protein expression, measured by luciferase enzyme activity assays. Negative regulation by the PUMs correlates with a reduction in reporter mRNA level, as measured by RT-qPCR and Northern blotting (8, 9). Importantly, mutations in the PRE that inactivate PUM1 and PUM2 binding completely alleviate

regulation of the reporter (8, 9). This powerful system provides a means of interrogating the mechanism of repression, and allowed us to discover the role of specific RNA degrading enzymes (8, 9). A similar approach can be taken to study other RNA regulatory sequences and/or trans-acting factors including activators or repressors.

Description of methods

Overview

Here we outline the steps needed to optimize the reporter-based model for studying RNA regulatory sequences including design of the reporter construct, transfection optimization, and measurement of reporter protein and RNA levels (Fig. 2.1). We discuss critical controls, data analysis, and keys to success. We provide protocols and examples resulting from our studies of mRNA regulation by PUF proteins (8, 9).

Reporter construct design

When designing reporter gene constructs, there are several key considerations: reporter of choice, promoter, 3' UTR, and transfection strategy. Numerous reporter genes are available, the most popular of which are luciferases (e.g. derived from firefly and *Renilla*) and fluorescent proteins (e.g. green fluorescent protein, GFP). Luciferase assays are sensitive and have a broad dynamic range, and different luciferases with unique substrates are available to permit multiplexing. For example, in our research on PUF protein mediated mRNA regulation, we used a dual luciferase reporter approach: *Renilla* luciferase serves to monitor regulation and firefly luciferase serves as an internal control (Fig. 2.2). Luciferases are amenable to multi-well plate analysis and they report on relative differences between populations of cells. Fluorescent proteins are commonly analyzed on a per cell basis by high content screening or flow cytometry and reported as the percentage of cells expressing the reporter. Reporters that have been

optimized for expression in mammalian cells and engineered to minimize cryptic regulatory sequences are ideal to maximize expression changes specific to the test sequence of interest (10, 11). When choosing the reporter and its corresponding plasmid expression vector, it is important to ensure they do not already contain the regulatory sequence.

An important consideration is the promoter driving reporter expression. The chosen promoter should be active in the chosen cell type. The promoter strength can have an impact on the ability to detect regulation; if expression is too strong, the regulatory mechanism can be overwhelmed. If the promoter is too weak, the signal to noise ratio would be too low. Ideally, the promoter should be minimal in size to avoid introducing unnecessary variables that might confound analysis.

The reporter construct should contain a well-defined transcription unit, including a defined transcription start site, 5' UTR and 3'UTR. Some reporter genes also contain an efficiently spliced intron, which may be beneficial for optimal expression. For analysis of 3'UTR mediated regulation, a minimal 3' UTR with a strong, well-defined cleavage/polyadenylation site is important. Ideally the 3' UTR should contain one or more unique restriction sites into which the test regulatory sequence can be cloned. Mammalian cleavage/polyadenylation signals can be somewhat degenerate; therefore, the SV40 signal is a good choice for efficient production of a well-defined 3' end.

To create a regulated reporter gene, the regulatory sequence of interest must be cloned into the proper context, such as the 3'UTR. As a critical control, create an unregulated reporter. To do so, the regulatory sequence should be mutated to inactivate its activity. If inactivating mutations are not yet known, the original reporter gene lacking a regulatory sequence is an appropriate control. The inserted regulatory sequence may be a minimal binding site for a

sequence specific regulatory protein (e.g. the Pum Response Element) or larger RNA sequence (e.g. a full length 3'UTR). A shorter sequence is less likely to contain multiple regulatory elements that affect expression. Be aware that full length 3' UTRs may contain alternative cleavage/polyadenylation signals that have the potential to cause heterogeneous processing of the reporter's 3' end.

Specific features will also be required depending on the transfection strategy used. Include a selectable marker if you plan to create stable cell lines. Constructs for transient transfections may include a second reporter that serves as an internal transfection control. Alternatively, a second plasmid can be co-transfected as the control. Lentiviral regulatory elements and strategy may be necessary to overcome difficulties associated with transfecting primary cells.

Transfections

Lipid-based transfection reagents that offer high efficiency with minimal impact on cell health are ideal for introducing reporter plasmids into the majority of cell lines. Cell type-specific protocols are available either from the transfection reagent manufacturer or in the scientific literature. For best results, optimal transfection conditions should be determined empirically (12). The transfection reagent will dictate the optimization variables. Important variables include the lipid to DNA ratio, mass of DNA, and the exposure time.

During the optimization phase of establishing transfection conditions, a single reporter can be used, such as the unregulated reporter. This removes any experimental treatment variables and focuses on achieving the best transfection conditions possible. Monitor both reporter expression and viability to find the optimal balance of high reporter expression with minimal toxicity from the procedure. For cell types that are sensitive to the transfection and

exhibit reduced cell health, incorporate a medium change after transfection using medium pre-equilibrated under optimal CO_2 and temperature conditions. Also, give consideration to the plasmid purification method used. Traditional purification chemistries can co-purify endotoxin (lipopolysaccharide from *E. coli* outer membrane) with the plasmid DNA that can cause toxicity or other inadvertent effects on cell biology. Choose a purification chemistry designed to give transfection-quality DNA with minimal endotoxin. Replicate transfections should always be performed to assess variability in the transfection technique. Highly variable transfections will negatively impact your ability to measure regulation. To assess and normalize sample to sample variation in transfection efficiency, it is advisable co-transfect an internal control. In the dual luciferase strategy, a plasmid expressing firefly luciferase is co-transfected with the *Renilla* luciferase reporter gene plasmid, and firefly luciferase activity is used to normalize the resulting data (see below).

Critical controls should be included in all transfection experiments. Untransfected and/or mock transfected cells are essential to measure background signal in the luciferase activity assays (and later, RNA assays). An unregulated reporter such as the reporter plasmid lacking regulatory sequences or containing mutated inactive regulatory sequences is crucial for discerning effects of a regulatory sequence of interest.

Cell culture considerations

There are several points to consider for optimizing performance; most are cell type-specific. The number of cells plated per well will depend on the plate format used, the cell line's optimal growth density (e.g. whether they thrive with contact or respond to contact inhibition), the growth rate, and intended length of the assay. The latter two points should be considered together, as the goal is to prevent overgrowth and exhaustion of the medium throughout the

duration of the assay. These effects will negatively impact cell health and thus the reliability of reporter assays.

Both the passage number and the confluence of the cells prior to transfections can affect the reproducibility of cell based assays. Cells generally take several passages after thawing to exhibit regular healthy growth patterns. Using them too soon will make them more likely to experience toxicity during the transfection. Later passage cells are susceptible to genetic changes that could alter the biological response of interest. Be cautious until you have determined the reliable passage number range in which the cells' responses are stable. The confluence prior to plating for transfection can also affect overall reproducibly: allowing cells to overgrow just prior to transfection can drive them into a senescent state, making them less amenable to transfection and exogenous gene expression. Monitor confluence closely and passage at a consistent density for optimal results.

The conditions during the plating process are another important consideration for optimal results. For adherent cells, it is important to allow a recovery period after trypsinization during which surface proteins recuperate and cells adhere to culture plates. For 96 well plate formats, the outer perimeter wells should be filled with medium, but no cells, to avoid causing detrimental edge effects. Edge effects are caused by evaporation from wells in the outer edges of the plate and may produce artifacts. The choice of plate should be considered carefully: luminescent cell-based assays should employ white tissue culture treated plates to maximize reflection and signal intensity. This is in contrast to fluorescent assays, which are typically performed in black plates to minimize background signal. Another key consideration is that signal bleed through between wells may be problematic if there is a wide range of signal intensity generated within a single

plate, thus, it is important to note that plates vary widely and that some may perform better than others.

Balance of multiple plasmids

It is often desirable and advisable to incorporate multiple plasmids into the transfection. When doing so, it is important to maintain the optimal transfection conditions. From there, it is simply a matter of adjusting the ratio of the reporter plasmids to find the best balance for your assay.

The most common reason to transfect multiple plasmids is to include an internal control to normalize for variation in transfection efficiency of the experimental reporter. The internal control should be constitutively expressed and, ideally, of the same vector backbone as the experimental reporter (for example, see Figure 2.2). Include the internal control as a minimal fraction of the total DNA transfected that still gives significant signal above background. To determine this, compare the signal from untransfected or mock transfected cells to that of cells transfected with increasing ratios of control to experimental plasmid. The signal from individual reporter plasmids and the internal control should be at least 3 standard deviations above the background signal. Typical mass ratios of experimental to internal control plasmid are in the range of 2:1 and 20:1.

An effector plasmid may be included to express a putative regulator, such as a trans-acting RNA binding protein. Effector plasmids are typically the highest proportion in the total transfected plasmid pool. Luciferase-encoding plasmids optimized for bright, efficient expression, can be included at relatively low amounts while still maintaining high signals above background, making them ideal for multi-plasmid transfections. Determine the minimum amount of reporter plasmid that produces significant signal above background. If studying a

repressor, it is important to leave enough range above background to accurately detect decreases in reporter regulation. The remainder of the plasmid pool may then be used to titrate in the effector plasmid. The total amount of plasmid DNA must be balanced to maintain the optimal mass of transfected DNA. When necessary, use an empty plasmid vector to balance the total mass. Ideally the empty vector is identical to the effector plasmid except that it lacks a protein coding region. In any experiment testing the impact of an effector plasmid, always include the following critical control: co-transfect the plasmid lacking the effector protein coding region. Alternatively, co-transfect a mutated inactive form of the effector; however, be aware that over-expressed mutant proteins may potentially have a dominant negative effect.

Protocol: Transfection of exogenous luciferase reporters

Here we describe transfection conditions in 96 well format for subsequent analysis of luciferase protein expression. The protocol is based on our work analyzing PUF mediated mRNA regulation by human PUM1 and PUM2 in human HEK293 cells (8). Transfections for RNA analysis can be performed as described below, but should be scaled up to yield sufficient amount of RNA. For instance, a 6 well dish format will produce 10-20 µg of cytoplasmic RNA.

1. Select mammalian reporter vector. We chose psiCHECK™-1 (Promega) which contains key features shown in Figure 2.2, including a constitutive SV40 promoter, the coding sequence for *Renilla* luciferase, a minimal 3'UTR with unique restriction sites, and an efficient cleavage/polyadenylation signal. To measure regulation by human PUFs, PUM1 and PUM2, 3 copies of the Pumilio Regulatory Element (PRE) were inserted into the 3' UTR of to produce psiCHECK-1 3xPRE (Fig. 2.2). As unregulated reporters, we use psiCHECK-1 No PRE, which lacks a PRE element, or psiCHECK-1 3xPREmt, which containing 3 mutant

PREs. As an internal control, we used the plasmid pGL4.13 (Promega) that encodes firefly luciferase with expression driven by the SV40 promoter (Fig. 2.2). To express effector proteins, we use the mammalian pFN21A vector (Promega) as it contains a strong CMV promoter to drive effector protein expression as a protein fusion to Halotag, which can be specifically detected by fluorescent labeling.

2. Purify plasmids using the PureYield™ Plasmid Midiprep System (Promega) or similar chemistry yielding transfection-quality plasmid DNA.
3. Measure the concentration and purity of the DNA using a spectrophotometer.
4. Culture HEK293 cells at 37°C under 5% CO_2 in DMEM and 10% FBS. Passage at 70-90% confluence. Use cells between 5-40 passages.
5. For reporter protein assays, plate cells in sterile, tissue culture-treated, white walled 96 well plates at 2 x 10^4 cells per 100µl per well. Allow the cells to recover overnight.
6. Transfect cells with FuGENE® HD (Promega) following the manufacturer's protocol. Use a 3:1 ratio of FuGENE HD:DNA. Allow DNA to complex with FuGENE HD for at least 10 but not more than 15 minutes at room temperature. Add a total mass of 100ng DNA per well of 96 well plate. Here are two examples of multiple reporter transfections:
    a. For dual reporter transfections without effector, transfect 75ng psiCHECK-1 reporter with 25ng pGL4.13 internal control per well.
    b. For dual reporter transfections with effector, transfect 10ng psiCHECK-1 reporter with 5ng pGL4.13 and a total of 85ng pFN21A-based effector expression plasmid per well. To titrate in the effector, divide the 85ng mass of pFN21A between effector-containing pFN21A (0-85ng) and pFN21A control (85-0ng, proportionately). In this case, we used pFN21A empty vector as the negative control.

7. Return cells to incubator for 48 hours.

Reporter protein analysis

Reporter protein assay

Luciferase activity assays are a convenient, rapid way to measure luciferase reporter protein expression. Light production, catalyzed by luciferase enzyme activity, is measured using a luminometer. Depending on the reaction chemistry, luciferase activity can be measured in a reaction with brief, high light output, known as flash format, or in a format that produces stabilized light output but reduced brightness. Use of flash reagents requires use of substrate injectors built into the detection instrument for immediate readings; while use of a stabilized light output is more amenable to multiwell plates. Assays can be performed by adding substrate to cell lysates or directly to cells in culture (i.e. homogeneous assays). The latter option reduces manipulation and hands-on time and is highly recommended for 96 well analyses. We use a dual luciferase assay format in which both *Renilla* reporter and firefly internal control can be sequentially measured in the same sample. Ensure all reagents and plates have been pre-equilibrated to room temperature before performing the assay.

The optimal time after transfection necessary to detect reporter activity is another important consideration. For early optimization experiments, defaulting to a convenient time ~24 hours after transfection is often sufficient. However, regulatory sequences will influence the peak expression time frame. Before experimental conditions are finalized, analyze replicate transfections at a range of time points to determine the time that gives the best measure of the experimental effect. Typically an optimal time is 18-48 hours after transfection. Avoid prolonged incubations in which cells might become overgrown, deplete their medium, and die.

Reporter assay data analysis

Analysis of reporter activity data is a multistep process and, if done diligently, will allow monitoring of experimental design and execution quality. This gives the means to quickly identify and troubleshoot issues that can arise. There are 3 phases to analyzing the data:

1. Evaluate raw data (luciferase data are expressed as Relative Light Units, RLU):
    a. Calculate the mean background. Background refers to the luciferase signal from untransfected or mock transfected cells. In dual reporter experiments, you will have a mean background for firefly and *Renilla* luciferases.
    b. Subtract the mean background from the corresponding RLU values from each transfected sample.

        $\text{Background subtracted RLU}_{\text{firefly}} = \text{RLU}_{\text{firefly}} - \text{mean Background}_{\text{firefly}}$

        $\text{Background subtracted RLU}_{\textit{Renilla}} = \text{RLU}_{\textit{Renilla}} - \text{mean Background}_{\textit{Renilla}}$
    c. Calculate the mean signal from replicates of background subtracted firefly and *Renilla* data and assess variability (Fig. 2.3A and 2.3B)

High variability between samples can indicate inconsistency in plating and/or transfections. When possible, use master mixes of common components. Strive for accuracy and precision in pipetting technique by using calibrated multichannel pipettes, changing tips between replicate pipetting, and ensuring even cell suspension during plating.

The internal control should remain relatively consistent across all transfected samples and be at least 3 standard deviations above background. The test reporter may fluctuate depending on the effect of the experimental variable. Variable internal control data may indicate variations in transfection efficiency. Normalization of your reporter to the internal control in the next stage of data analysis will allow you to correct for this variation (13). In some instances an effector

may alter internal control expression. A strategy to compensate for this is to use the same vector backbone for both the reporter and internal control, varying only the reporter and the test regulatory sequences (13). In this way, the internal control will normalize the effects that are independent of the test regulatory sequences.

2. Normalize reporter data by calculating the Relative Response Ratio (RRR):
    a. For each sample, divide the reporter luciferase value by the corresponding internal control luciferase value to yield a Relative Response Ratio (13). Use background subtracted RLU for this calculation. The luciferases used for the reporter versus the internal control will depend on your experimental design.

        $RRR_{regulated} = RLU_{regulated\ reporter} / RLU_{internal\ control}$

        $RRR_{unregulated} = RLU_{unregulated\ reporter} / RLU_{internal\ control}$

    b. Calculate the mean of all replicate RRR and assess variability (Fig. 2.3C).

3. Calculate change for all samples to determine the effect of the regulatory sequence on protein levels
    a. % repression = $[1 - (RRR_{regulated} / \text{mean } RRR_{unregulated})] \times 100$,

        or for activators, calculate fold change = $(RRR_{regulated} / \text{mean } RRR_{unregulated})$
    b. Determine the mean of the replicates and assess variability (Fig. 2.3D).

Protocol: Reporter activity analysis using dual reporter assay

The following protocol is based on our work analyzing PUF mediated mRNA regulation by human PUM1 and PUM2 in human cells and is a continuation of the protocol outlined above (8).

1. Forty-eight hours after transfection, measure luciferase activity with the Dual-Glo® Luciferase Assay System (Promega) using a GloMax®-Multi+ plate-reading luminometer (Promega) according to the manufacturer's instructions.
2. Analyze the data:
    a. Average the firefly and *Renilla* RLUs from the mock transfected wells. This is background.
    b. Subtract the average background from the corresponding luciferase RLUs in each individual well.
    c. Calculate the mean *Renilla* RLU and standard error of the mean (SEM) for each background-subtracted sample set. Do the same for the firefly luciferase RLU values.
    d. Normalize data by calculating the Relative Response Ratio (RRR). Using the background-subtracted values for each sample well, divide the experimental reporter value (*Renilla*) by the corresponding internal control value (firefly activity from pGL4.13). Calculate the mean RRR and Standard Error of the Mean (SEM) for each sample set and graph the resulting data. As an example, Figure 2.4 shows RRR values from three reporters that we used to measure PUM Response Element mediated repression of *Renilla* luciferase expression in HEK293 cells. Reporters, described in Figure 2.2, include the unregulated psiCHECK-1 No PRE reporter and the regulated reporter, psiCHECK-1 3xPRE, which is bound and repressed by endogenous PUF proteins, PUM1 and PUM2. The unregulated reporter psiCHECK-1 3xPREmt, wherein the PRE was mutated to prevent PUF binding was also tested (Fig. 2.4). In this example, the presence of wild type PREs cause a substantial

reduction in the RRR values, relative to the unregulated No PRE reporter (Fig. 2.4). Mutation of the PREs alleviated the repressive effect (Fig. 2.4).

e. To determine the impact of an effector protein on the regulation, RRR values can be calculated for psiCHECK-1 No PRE, psiCHECK-1 3xPRE, and psiCHECK-1 3xPREmt reporters at each mass of transfected effector plasmid (Fig. 2.4, effector plasmid amounts of 0, 20, 50, or 85ng). Recall that the total mass of transfected DNA in each sample is held constant by balancing with expression vector lacking the effector protein coding sequence.

f. Calculate the percent repression caused by the regulatory element. For PUF mediated repression via the 3xPRE (8, 9), the equation % repression = [1 – ($RRR_{psiCHECK\text{-}1\ 3xPRE}$ /mean $RRR_{psiCHECK\text{-}1\ no\ PRE}$)] x 100 was used. $RRR_{psiCHECK\text{-}1\ 3xPRE}$ are from the samples transfected with psiCHECK-1 3xPRE and pGL4.13. Mean $RRR_{psiCHECK\text{-}1\ no\ PRE}$ is from the samples transfected with psiCHECK-1 No PRE and pGL4.13. The resulting percent repression data are shown in Figure 2.4.

g. To ascertain the change in PRE mediated repression caused by the effector, percent repression values are determined for psiCHECK-1 3xPRE relative to unregulated psiCHECK-1 No PRE and compared to the values from samples transfected with empty effector plasmid (Fig. 2.4, effector = 0 ng). In the example shown in Figure 2.4, the effector was a dominant negative form of the deadenylase enzyme CNOT8, which specifically reduced percent repression of the 3xPRE reporter in dose dependent manner (8, 9)

Reporter RNA analysis

RNA purification

Mammalian RNA should be purified quickly after harvesting cells in an RNase free environment, using RNase free reagents and equipment. Aerosol barrier tips should be used for all pipetting steps. A number of resources are available for more information on the maintenance of RNase free conditions (14, 15). Multiple methods are available for purifying RNA. Most techniques yield sufficiently high quality RNA for this application, but some purification systems offer advantages through streamlined methods that minimize potential error, save time, and allow batch processing. Regardless of the purification protocol used, it is always best to purify RNA from fresh cells. Trizol extraction uses phenol and guanidinium isothiocyanate to isolate RNA and is an inexpensive means of RNA purification. Kits for RNA purification are readily available from a variety of manufacturers and use column purification to isolate RNA. Keep in mind that the binding capacity of columns in each kit varies. Column purification kits and automated nucleic acid purification systems are fast and reliable, and will aid in processing multiple samples in a short period of time to consistently yield high quality RNA. Our laboratory uses a modification of the automated Maxwell® 16 LEV simplyRNA purification system (Promega), which allows us to purify high quality RNA samples in about an hour.

For RNA extractions, scale up transfections to larger well dishes depending on the RNA yield desired. Although RNA can be extracted from the same 96 well format used above, higher yields are typically needed for RNA analysis, especially during optimization. Scale up by adjusting all components of the transfection proportionately to the differences in well surface area.

Contaminating plasmid DNA present in RNA preparations is often the main culprit contributing to high background signal in qPCR assays. A number of methods can be used to reduce such background in RNA measurements. DNA can be reduced by isolating the cytoplasmic pool of RNA. This involves mild detergent-based cell lysis and removal of nuclei by centrifugation prior to extracting the RNA. As this is not sufficient to remove all plasmid DNA, additional DNase treatment is necessary (discussed in next section).

Once an RNA sample has been obtained, perform quality control measurements to assess its purity and integrity before proceeding with analysis of gene expression. The UV absorbance of RNA should be measured to determine concentration and purity. A 260nm/280nm ratio of 1.8-2.2 indicates sufficient purity. Next, to check integrity of the RNA, perform formaldehyde denaturing agarose gel electrophoresis on 0.5-1 µg of RNA (14). The 18S and 28S ribosomal bands on the gel should be clear and crisp, not smeared, and the 28S band should be roughly twice the intensity of the 18S. Alternatively, a Bioanalyzer (Agilent) can be used to assess RNA integrity.

Removal of contaminating DNA using DNase enzyme

Plasmid DNA contamination in RNA samples can mask changes in gene expression detected by qPCR; therefore, it is crucial to treat RNA samples extensively with RNase-free DNase enzyme. Standard DNase treatments included during some purification protocols can be effective at minimizing genomic DNA; however, plasmid DNA may persist after purification. Therefore, an additional extensive DNase treatment after RNA purification should be included. Importantly, divalent metal ions necessary for DNase activity pose a hazard to the integrity of the RNA samples: heating in the presence of these divalent metals hydrolyzes RNA. Therefore, do

not heat inactivate DNase. Instead, precipitate the RNA with ethanol and resuspend in a solution of EDTA and EGTA chelators to protect the RNA during subsequent heating steps.

Reverse transcription and quantitative PCR analysis of reporter gene expression

Reverse transcription is performed to create complementary DNA from the RNA. Reverse transcriptases require magnesium. Depending on the volume of RNA used in the RT reaction it may be necessary to adjust the magnesium concentration to offset the presence of chelators. Reverse transcriptases also require a primer to initiate transcription. This primer can be gene specific for detection of a single RNA, oligo-dT for detection of poly-adenylated mRNAs, or random hexamers for detection of all RNAs. All steps are performed using RNase free aerosol barrier pipette tips.

Once cDNA has been prepared, sequences are detected using gene-specific primers in quantitative PCR. Several qPCR chemistry choices exist based on the use of probes, fluorescent primers or DNA binding dyes (16-19). Variability in qPCR measurements can be assessed by performing replicates including technical and biological replicates. To minimize variability, use master mixes and calibrated multichannel pipettes when possible. Also, always use low retention aerosol barrier pipette tips.

Critical controls must be performed to interpret the data with confidence. First, a control PCR is performed in the absence of nucleic acid. This “no template control (NTC)” measures environmental contamination. Another critical control is performed in the absence of reverse transcriptase but in the presence of RNA template. qPCR is then performed on these “no RT control” samples to assess the presence of contaminating DNA in the RNA preparation or identify nonspecific products generated by the qPCR primers.

Quantitative PCR primer design considerations

There are many resources widely available to aid in the design and optimization of qPCR assays (16, 18-20). We suggest consulting the technical manual of the qPCR kit you plan to use for specific optimization parameters. The design of efficient and specific qPCR primer sets is crucial for detecting changes in gene expression. There are a number of helpful resources and software programs available to assist in primer design such as Primer 3 and Lasergene (DNASTAR) (20). The primary goal of qPCR primer design is to create primers that amplify a specific small target, generally between 75 and 250 nucleotides in length. In general, primers should be 17-25 nucleotides in length and contain between 50 to 60 percent GC content. Primer 3' ends should not be complementary to avoid hairpin or dimer formation. We typically aim to create primers with melting temperatures near 60°C. Commercial primer sets are also available but may require optimization. Perform a BLAST search using the primer sequences against the relevant genome database (NCBI) and the reporter plasmid sequences to confirm specificity of the primers for target sequences. New primer sets should be validated by confirming that they amplify the expected PCR product, as determined by DNA agarose gel electrophoresis. For each new primer set, qPCR conditions are optimized for primer concentration and annealing temperature. A standard curve is used to evaluate amplification efficiency (19, 20). Use conditions under which the primers are most efficient (between 90 and 110%) for subsequent qPCR reactions. Multiple replicates should be performed to assess experimental variability.

qPCR data analysis

Reporters and internal control mRNAs can be measured in separate qPCR reactions. Alternatively, both can be measured simultaneously in the same sample using a multiplexed qPCR assay. Multiplexed qPCR relies on measurement of fluorescence from primers containing

different fluorophores. For either format, qPCR raw data are expressed as cycle threshold (Ct) values. For each sample, the Ct value of the reporter is normalized to the Ct value of the internal control, thereby normalizing for variation in transfection efficiency. To do so, calculate the ΔCts as follows (14-16, 18-20):

$$\Delta Ct_{regulated} = Ct_{regulated\ reporter} - Ct_{internal\ control}$$

$$\Delta Ct_{unregulated} = Ct_{unregulated\ reporter} - Ct_{internal\ control}$$

To assess the standard error, calculate both the regulated and the unregulated results relative to the mean value of the unregulated ΔCt. Use the comparative ΔΔCt method to calculate ΔΔCt for each sample (16-20):

$$\Delta\Delta Ct_{regulated} = \Delta Ct_{regulated} - \text{mean } \Delta Ct_{unregulated}$$

$$\Delta\Delta Ct_{unregulated} = \Delta Ct_{unregulated} - \text{mean } \Delta Ct_{unregulated}$$

Finally, for all samples calculate the relative change to determine the effect of the regulatory sequence on mRNA levels:

Percent repression = $100*[1 - 2^{-\Delta\Delta Ct}]$, or for activation,

calculate fold change = $2^{-\Delta\Delta Ct}$

Determine the mean repression or change for each sample and assess variability. See Figure 2.5 for an example of these calculations.

Protocols

Purification of human cytoplasmic RNA from HEK293 cells

1. Harvest cells 48 hours after transfection: aspirate growth medium, rinse cells once with sterile, room temperature 1X PBS, transfer plate to ice and add 175 µl/well ice cold cytoplasmic RNA lysis buffer (50 mM Tris-HCl, pH 8.0, 140 mM NaCl, 1.5 mM Mg $MgCl_2$,

0.5% Nonidet P-40 and 1000 units/ml RNasin® Ribonuclease Inhibitor (Promega) added immediately prior to use). Scrape cells with a cell scraper and transfer lysates to pre-chilled 1.5 mL microfuge tubes on ice.

2. Incubate lysates on ice for 10 minutes then centrifuge lysates at 300 *x g* for 5 minutes at 4°C.
3. Carefully transfer supernatant to fresh pre-chilled 1.5 mL microfuge tube and keep on ice.
4. Prepare Maxwell® 16 LEV simplyRNA Blood cartridges according to manufacturer instructions.
5. Add 200 µL ice cold simplyRNA Homogenization Solution (with 1-thioglycerol added) to lysates and transfer tubes to room temperature. Mix completely by vortexing.
6. Add 200 µL simplyRNA Lysis Buffer to lysates. Mix well by pipetting.
7. Transfer lysates to well #1 of Maxwell® 16 LEV simplyRNA cartridge, transfer cartridge rack to Maxwell® 16 instrument, and run the Maxwell simplyRNA Blood purification program.
8. Measure RNA concentration and purity by UV absorbance. Analyze 0.5-1 µg RNA by denaturing agarose gel electrophoresis.

DNase treatment of cytoplasmic RNA samples with DNase

1. Dilute 10 µg RNA to a final volume of 38 µL with nuclease-free water.
2. Add 5 µL 10X Turbo™ DNase (Life Technologies) reaction buffer to the RNA sample.
3. Add 2 µL RNasin® Ribonuclease Inhibitor.
4. Add 5 µL Turbo™ DNase (2 units/µL). We suggest using 1 unit per µg of RNA, which is in excess of the manufacturer instructions but worked well for minimizing plasmid DNA background in qPCR.
5. Mix well by pipetting and incubate at 37°C for 3 hours.

At the end of the incubation, **DO NOT** heat inactivate the DNase enzyme as this can cause divalent metal-mediated RNA hydrolysis. Proceed directly with ethanol precipitation.

Ethanol precipitation of RNA

1. Add 10% v/v 3M sodium acetate pH 5.2 to DNase reaction.
2. Add 2.5X reaction volume of 100% ethanol and mix well. Glycogen may be added as a carrier for small amounts of RNA.
3. Precipitate RNA at -80°C or on dry ice for at least 1 hour.
4. Centrifuge samples at 14,000 x g in a 4°C microfuge for 30 minutes.
5. Carefully aspirate the supernatant without disrupting the RNA pellet.
6. Add 900 µL 70% ethanol to each RNA sample and vortex well for 30 seconds.
7. Centrifuge samples at 14,000 x g in a room temperature microfuge for 15 minutes.
8. Aspirate supernatant without disrupting pellet.
9. Briefly spin samples to collect remaining ethanol.
10. Use a P-10 pipette tip to wick away residual ethanol by capillary action.
11. Uncap samples and let sit for 5 minutes on the bench to allow remaining ethanol to evaporate. Do not dry more than 5 minutes, as it will become very difficult to resuspend the RNA pellet.
12. Resuspend pellet in 10 µL sterile 10 mM EDTA, 10 mM EGTA. Vortex and incubate briefly at 37°C to dissolve the pellet.
13. Assess RNA concentration by UV absorbance using a spectrophotometer, blanking the instrument with the resuspension solution.
14. When optimizing RNA purification protocols, we recommend visualizing the RNA samples using denaturing gel electrophoresis after DNase treatment and ethanol precipitation.

Reverse transcription

1. Dilute 1 µg RNA to 250 ng/µL in water. Assuming the initial stock of RNA is ~1µg/µl, the final concentration of EDTA and EGTA in RT reactions will be 0.5 mM each which is sufficiently low to not interfere with the GoScript™ Reverse Transcriptase (Promega) reaction**.**
2. Add 4 µL RNA to 1 µL random hexamers (500 ng/µL). Include a no RT control reaction for each sample.
3. Incubate at 65°C for 5 minutes. Transfer to ice for 5 minutes. Centrifuge briefly to collect any evaporated liquid in the lids of tubes.
4. Prepare reverse transcription master mixes and proceed with the reverse transcription reaction according to the manufacturer instructions. Use 3 mM $MgCl_2$ and 20 units RNasin per final 20 µl total reaction.
5. Proceed immediately to qPCR or store cDNA at -20°C until ready to use.

Multiplexed quantitative PCR detection of luciferase cDNAs

1. Using the Plexor® qPCR System (Promega), have the following primers synthesized (Biosearch Technologies) and prepare 2.5 µM dilutions of each primer set in Plexor® qPCR System MOPS/EDTA Buffer.
   - firefly forward: 5'-dGATCCTCAACGTGCAAAAGAAGC- 3'
   - firefly reverse: 5'-d FAM-isoC-TCACGAAGGTGTACATGCTTTGG-3'
   - *Renilla* forward: 5'-d CAL Fluor Orange 560-isoC-CGCAACTACAACGCCTACCTTC-3'
   - *Renilla* reverse: 5'-dCCCTCGACAATAGCGTTGGAAAA-3'

2. Thaw all Plexor® reagents on ice.

3. Prepare qPCR master mix as indicated below, scaling up for the number of samples plus an additional 10%.

Component	**Volume (μL) per amplification**
2X Plexor Master Mix	25
Renilla primer mix (2.5 μM ea)	1
firefly primer mix (2.5 μM ea)	1
Water	18
Total volume	45

4. Pipet 5 μL reverse transcription reaction per well of a 96 well qPCR plate. We typically perform duplicate qPCR reactions on each cDNA sample to monitor technical variability. Include corresponding no RT and no template controls.
5. Add 45 μL PCR master mix to each well and pipet to mix.
6. Seal the plate thoroughly using adhesive qPCR film and centrifuge briefly to collect liquid to the bottom of the wells.
7. Amplify using the following thermal cycling conditions. We used the CFX96™ Real-Time PCR Detection System (Bio-Rad):
    a. 95°C 2 minutes
    b. 95°C 5 seconds
    c. 60°C for 35 seconds
    d. Repeat steps b-c 40 times

e. Perform thermal melting curve

8. Upon completion of the program, export the qPCR data from the CFX Manager Software and import it into the Plexor® Analysis Software (Promega).

9. Analyze the data (As an example, see Figure 2.5):

   a. For each regulated reporter sample, calculate $\Delta Ct_{psiCHECK\text{-}1\ 3xPRE} = Ct_{psiCHECK\text{-}1\ 3xPRE} - Ct_{pGL4.13}$

   b. For each unregulated reporter sample, calculate $\Delta Ct_{psiCHECK\text{-}1\ No\ PRE} = Ct_{psiCHECK\text{-}1\ No\ PRE} - Ct_{pGL4.13}$

   c. Then calculate the mean $\Delta Ct_{psiCHECK\text{-}1\ No\ PRE}$

   d. Next, for each regulated reporter sample, calculate $\Delta\Delta Ct_{psiCHECK\text{-}1\ 3xPRE} = \Delta Ct_{psiCHECK\text{-}1\ 3xPRE} - \text{mean } \Delta Ct_{psiCHECK\text{-}1\ No\ PRE}$

   e. For each unregulated reporter sample, calculate $\Delta\Delta Ct_{psiCHECK\text{-}1\ No\ PRE} = \Delta Ct_{psiCHECK\text{-}1\ No\ PRE} - \text{mean } \Delta Ct_{psiCHECK\text{-}1\ No\ PRE}$

   f. Finally, for each regulated reporter sample, calculate

      $\text{Percent Repression}_{psiCHECK\text{-}1\ 3xPRE} = 100*[1 - 2\text{^-}\Delta\Delta Ct_{psiCHECK\text{-}1\ 3xPRE}]$

   g. For each unregulated reporter sample, calculate

      $\text{Percent repression}_{psiCHECK\text{-}1\ No\ PRE} = 100*[1 - 2\text{^-}\Delta\Delta Ct_{psiCheck\text{-}1\ No\ PRE}]$

   h. Calculate mean % repression and assess variability for all replicate samples. The data shown in Figure 2.5 demonstrate that the presence of the PREs caused a 78% decrease in the amount of reporter mRNA from psiCHECK-1 3xPRE, measured relative to the unregulated psiCHECK-1 No PRE reporter. This result is consistent with the role of PUF proteins in binding to PRE containing mRNAs and promoting their degradation (8).

Concluding remarks

The methods described here may be adapted to study different aspects of post-transcriptional mRNA regulation including: mRNA processing, mRNA degradation including deadenylation, decapping, decay by exo- and endo-nucleases, and translational control. These approaches are widely applicable to study of trans-acting RNA binding proteins and small regulatory RNAs, such as microRNAs and short interfering RNAs. Once an assay is established that specifically measures regulation, rapid progress can be made through structure/function and mutational analysis of the regulator sequences and effectors. Furthermore, RNA interference mediated depletion, genetics, and/or cotransfection of regulators can be used to interrogate the factors and pathways involved in the mechanism of regulation. Reporter genes offer versatility in testing experimental parameters using quantitative assays, but it is always essential to validate data from reporter gene systems by returning the biological context of natural endogenous mRNAs.

References

1. Moore, M. J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309**, 1514–1518
2. Garneau, N. L., Wilusz, J., and Wilusz, C. J. (2007) The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8**, 113–126
3. Wickens, M., Bernstein, D. S., Kimble, J., and Parker, R. (2002) A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet.* **18**, 150–157
4. Wang, X., Zamore, P. D., and Hall, T. M. (2001) Crystal structure of a Pumilio homology domain. *Molecular Cell* **7**, 855–865
5. Cheong, C.-G., and Hall, T. M. T. (2006) Engineering RNA sequence specificity of Pumilio repeats. *Proc Natl Acad Sci USA* **103**, 13635–13639
6. Jenkins, H. T., Baker-Wilding, R., and Edwards, T. A. (2009) Structure and RNA binding of the mouse Pumilio-2 Puf domain. *JOURNAL OF STRUCTURAL BIOLOGY* **167**, 271–276
7. Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., and Gerber, A. P. (2008) Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System. *PLoS ONE* **3**, e3164
8. Van Etten, J., Schagat, T. L., Hrit, J., Weidmann, C., Brumbaugh, J., Coon, J. J., and Goldstrohm, A. C. (2012) Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *Journal of Biological Chemistry*
9. Weidmann, C. A., and Goldstrohm, A. C. (2012) Drosophila Pumilio protein contains multiple autonomous repression domains that regulate mRNAs independently of Nanos and brain tumor. *Molecular and Cellular Biology* **32**, 527–540
10. Fan, F., and Wood, K. V. (2007) Bioluminescent assays for high-throughput screening. *Assay Drug Dev Technol* **5**, 127–136
11. Paguio, A., Almond, B., Fan, F., Stecha, P., Garvin, D., Wood, M., and Wood, K. (2005) Luciferase Reporters: Less is More. *Promega Notes*, 7–10
12. Schagat, T., and Kopish, K. (2010) Optimize Transfection of Cultured Cells. *Promega Corporation Web site*, 1–5 [online] http://www.promega.com/resources/articles/pubhub/optimize-transfection-of-cultured-cells (Accessed February 18, 2013).
13. Schagat, T., Paguio, A., and Kopish, K. (2007) Normalizing Genetic Reporter Assays: Approaches and Considerations for Increasing Consistency and Statistical Significance. *Cell Notes*, 9–12 [online] http://www.promega.com/resources/articles/pubhub/cellnotes/normalizing-genetic-reporter-assays/).
14. Rio, D. C., Manuel Ares, J., and Timothy W Nilsen, P. D. (2010) *RNA*,
15. Sambrook, J., and Russell, D. D. W. (2001) *Molecular Cloning*, CSHL Press
16. Livak, K. J., and Schmittgen, T. D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408
17. VanGuilder, H. D., Vrana, K. E., and Freeman, W. M. (2008) Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques* **44**, 619–626
18. Schmittgen, T. D., and Livak, K. J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* **3**, 1101–1108
19. Pfaffl, M. W. (2001) A new mathematical model for relative quantification in real-time

RT-PCR. *Nucleic Acids Research* **29**, e45
20. Stratagene (2004) Introduction to Quantitative PCR: methods and applications guide. 1–50 [online] http://www.stratagene.com (Accessed February 12, 2013).

Figures

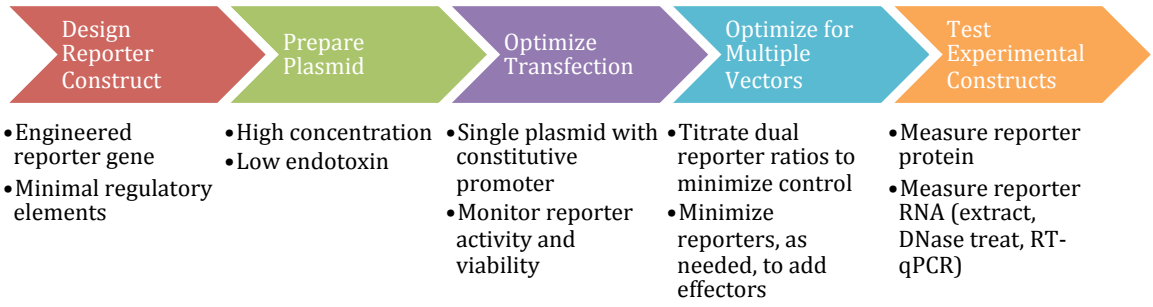


Figure 2.1 Workflow for using reporter genes to measure mRNA regulation in mammalian cells.

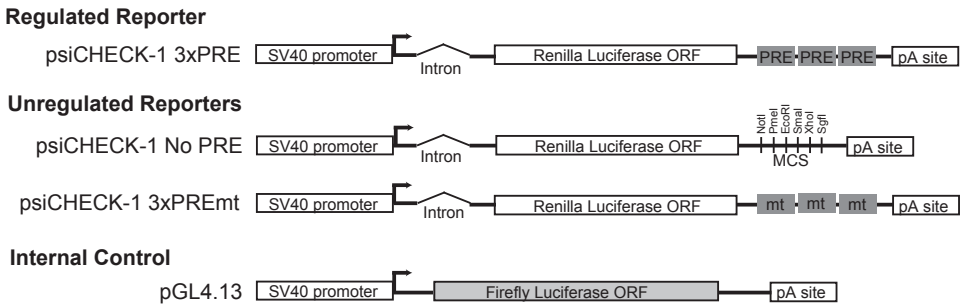


Figure 2.2 Schematic of regulated and unregulated reporter constructs used in cell based luciferase assays.

These specific reporters were created using the psiCHECK-1 plasmid (Promega). Reporter transcription is mediated by the SV40 promoter, derived from Simian Virus 40. An efficiently spliced intron is present in the 5'UTR to promote expression. The reporter encodes the *Renilla* luciferase open reading frame (ORF). The 3'UTR of psiCHECK-1 No PRE contains a multi-cloning site (MCS) with 6 unique restriction sites into which regulator elements can be inserted. Three copies of the Pumilio Response Element (PRE) or mutated PRE (mt) were inserted into the reporter 3'UTR. Cleavage and polyadenylation elements are labeled as "pA site."

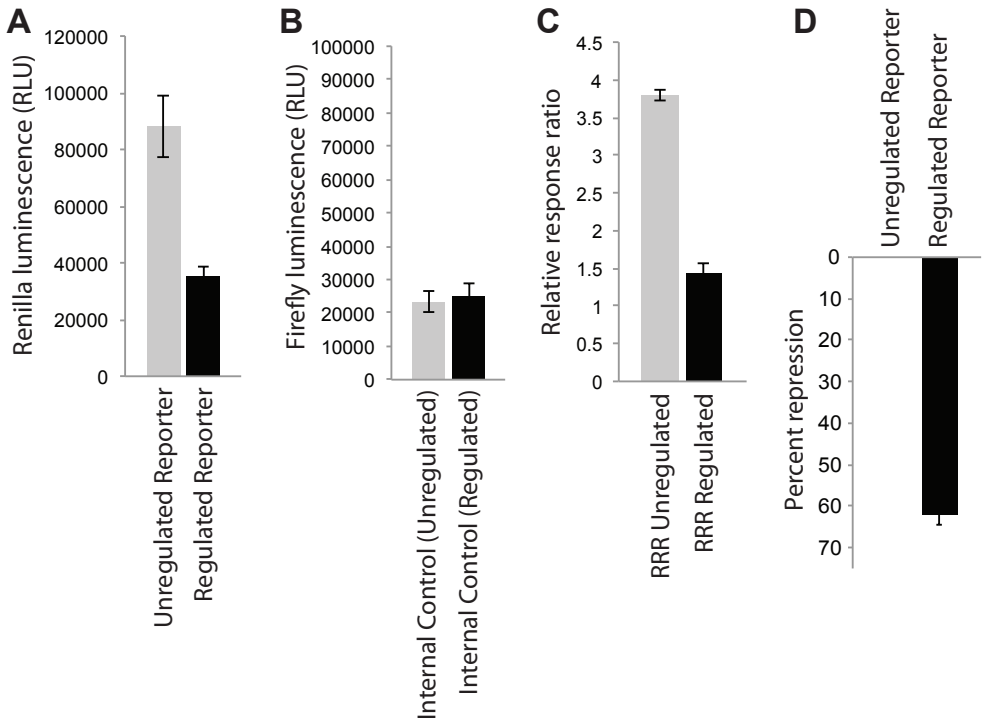


Figure 2.3 Example of dual luciferase assay data output.

A. Mean *Renilla* luminescence, measured as Relative Light Units (RLU) of regulated and unregulated reporters. B. Mean firefly luminescence from internal control that was co-transfected with either the regulated and unregulated *Renilla* luciferase reporters. C. Relative response ratio (RRR), or *Renilla* RLU/firefly RLU of both unregulated and regulated reporters. D. Percent repression of the regulated reporter relative to the unregulated reporter. These data are derived from our analysis of repression mediated by Pum Response Elements (PREs) in human cells. Regulated reporter psiCHECK-1 3xPRE encodes *Renilla* luciferase and contains three PREs whereas unregulated reporter psiCHECK-1 3xPREmt contains mutant PREs. The pGL4.13 plasmid, encoding firefly luciferase, served as the internal control. The data demonstrate that the wild type PRE elements cause repression.

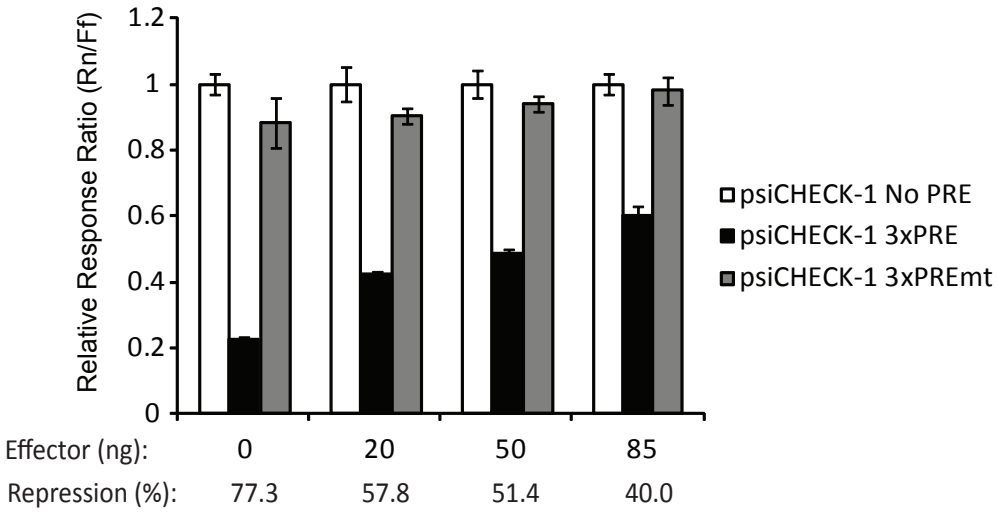


Figure 2.4 Example of overexpression of an effector protein and its effect on relative response ratios of regulated and unregulated reporters.

The amount of transfected effector plasmid per well of a 96 well plate is shown at the bottom. Effector protein mass was balanced using an identical plasmid that lacked the effector coding region. In this example, the effector protein was a mutant form of CNOT8, which acts in a dominant negative manner to block mRNA degradation and repression of the regulated reporter, psiCHECK-1 containing three PRE elements (psiCHECK-1 3xPRE), but not the unregulated reporters that lack a PRE (psiCHECK-1 No PRE) or contain mutant PREs (psiCHECK-1 3xPREmt). The percent repression in each condition is indicated at the bottom of the figure. In this experiment, percent repression for each test condition was calculated relative to the unregulated reporter, psiCHECK-1 No PRE. The data demonstrate that the effector inhibits PRE mediated repression.

Regulated Reporter: psiCHECK-1 3xPRE

	Internal Control Firefly Ct	Reporter Renilla Ct	ΔCt	Mean ΔCt	ΔΔCt	Fold Change	Mean Percent Repression	Standard Error
Sample 1	24.1	24.1	0	-0.1	2.33	0.20	79	1.0
Sample 2	24.3	24.1	-0.2		2.13	0.23		
Sample 3	24.6	24.4	-0.2		2.13	0.23		

Unregulated Reporter: psiCHECK-1 No PRE

	Internal Control Firefly Ct	Reporter Renilla Ct	ΔCt	Mean ΔCt	ΔΔCt	Fold Change	Mean Percent Repression	Standard Error
Sample 1	26.5	24.4	-2.1	-2.3	0.23	0.85	0.0	17
Sample 2	26.5	24.4	-2.1		0.23	0.85		
Sample 3	28.4	25.6	-2.8		-0.47	1.38		

Figure 2.5 Example of calculations of percent repression of reporter mRNA levels from data obtained using reverse transcription and multiplexed, quantitative PCR.

Renilla luciferase mRNA levels were measured in three samples expressing the regulated psiCHECK-1 3xPRE or the unregulated psiCHECK-1 No PRE reporters. In each sample, firefly luciferase mRNA levels from internal control pGL4.13 were also measured. Cycle threshold values (Ct) for each sample are reported. Calculations were then made as described in sections 2.5.5 and 2.5.6.5. Percent repression of psiCHECK-1 3xPRE mRNA level was then determined relative to psiCHECK-1 No PRE.

CHAPTER 3

Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs

This article was originally published in the Journal of Biological Chemistry. Van Etten, J., Schagat, T.L., Hrit, J., Weidmann, C.A., Brumbaugh, J., Coon, J., and Goldstrohm, A. Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *Journal of Biological Chemistry.* 2012, Vol 287: 36370-36383. Van Etten wrote the manuscript and carried out experiments in figures 3.1F, 3.1G, 3.3, 3.4A, 3.5, 3.6B-D, and 3.7. Schagat carried out experiments in figures 3.1A, C, D, and 3.E, 3.2, and prepared samples for mass spectrometry. Hrit carried out experiments in figures 3.1B, 3.4B-D. Weidmann carried out the control in figure 3.6A. Brumbaugh and Coon carried out mass spectrometry analysis.

Summary

PUF proteins are a conserved family of eukaryotic RNA binding proteins that regulate specific mRNAs: they control many processes including stem cell proliferation, fertility and memory formation. PUFs repress protein expression from their target mRNAs but the mechanism by which they do so remains unclear, especially for humans. Humans possess two

PUF proteins, PUM1 and PUM2, which exhibit similar RNA binding specificities. Here we report new insights into their regulatory activities and mechanisms of action. We developed functional assays to measure sequence specific repression by PUM1 and PUM2. Both robustly inhibit translation and promote mRNA degradation. Purified PUM complexes were found to contain subunits of the CCR4-NOT (CNOT) complex, which contains multiple enzymes that catalyze mRNA deadenylation. PUMs interact with the CNOT deadenylase subunits *in vitro* and *in vivo*. We used three approaches to determine the importance of deadenylases for PUM repression: First, dominant negative mutants of CNOT7 and CNOT8 reduced PUM repression. Second, RNA interference depletion of the deadenylases alleviated PUM repression. Third, the poly(A) tail was necessary for maximal PUM repression. These findings demonstrate a conserved mechanism of PUF mediated repression via direct recruitment of the CCR4-POP2-NOT deadenylase leading to translational inhibition and mRNA degradation. A second, deadenylation independent mechanism was revealed by the finding that PUMs repress an mRNA that lacks a poly(A) tail. Thus, human PUMs are repressors capable of deadenylation-dependent and independent modes of repression.

Introduction

Messenger RNAs (mRNAs) are subject to extensive regulation throughout their lifespan (1). Synthesis and processing events of precursor mRNAs in the nucleus are regulated to yield mature mRNAs. Once exported to the cytoplasm, translation and stability of mRNAs are controlled to ensure that the appropriate amount of encoded protein is produced at the proper time and cellular location. The discovery of factors and mechanisms responsible for gene regulation is crucial to deepening our understanding of how misregulation contributes to disease.

PUF (Pumilio and Fem-3 binding factor) proteins are trans-acting factors that regulate mRNAs by binding specific sequences in 3' untranslated regions (3'UTR)(2). Members of the PUF family share a conserved RNA binding domain composed of eight alpha helical repeats (3-8). These PUF repeats adopt a crescent shape, whose concave side binds to specific single-stranded RNA sequences. Each PUF repeat recognizes a single ribonucleotide base, mediated by three RNA recognition amino acids, and these contacts dictate the RNA binding specificity of each individual PUF protein (7).

Humans and other vertebrates possess two canonical PUF proteins, PUM1 and PUM2, collectively referred to as PUMs (9). PUMs share significant sequence similarity: amino acids outside of their RNA binding domains (RBD) share 75% identity whereas those within are 91% identical (9,10). Both PUM1 and PUM2 bind with high affinity to the consensus sequence UGUANAUA, hereon referred to as a PUM response element, PRE (7,11-13). PUMs are widely expressed in tissues and cell types (9,14). Given their similar RNA binding specificities and broad expression, it is possible that PUMs compete for many of the same mRNAs, supported by identification of mRNAs that associate with PUMs (13).

Genetics in model organisms demonstrated that PUFs control embryonic development, fertility, stem cell proliferation, and neurological functions, including the formation of memories (4,5,15-29). In mice, PUFs are involved in fertility through control of spermatogenesis (30,31). In cell culture, human PUMs were reported to affect cell proliferation (32). Insight into the functions of human PUMs emerged from identification of mRNAs that co-immunopurified with each PUM (13,33). Hundreds of mRNAs were enriched in the PUM1 and PUM2 immunoprecipitates, suggesting an extensive regulatory network. The list of putative target mRNAs included genes involved in gene expression, signal transduction, cell cycle and proliferation, among others (13,33). Importantly, regulation of these mRNAs remains to be demonstrated.

In model organisms, PUFs generally act as repressors, though in several instances they affect mRNA localization and perhaps activation of mRNAs (2,34-37). In *Drosophila* and *C. elegans,* PUF repression correlates with shortening of the poly(A) tail (i.e. deadenylation) (29,38-41). In yeast, PUFs repress by promoting degradation of target mRNAs, specifically deadenylation and decapping (42-46). Yeast PUFs bind to the Pop2 deadenylase subunit to enhance removal of the poly(A) tail, catalyzed by the Ccr4 deadenylase (43-45). The consequences of deadenylation are known to include translational down-regulation and initiation of mRNA degradation (47). In other cases, PUFs have been reported to directly inhibit translation, with several potential mechanisms having been proposed (5,48-51).

The mechanism(s) of mRNA regulation by human PUMs remains to be elucidated and a complete understanding of PUM repression will facilitate identification of biologically relevant target mRNAs. A repressive role for human PUMs is supported by several observations: Over-expression of PUM2 reduces expression of reporter genes (52) and over-expression of PUM

together with a putative partners NANOS3 was reported to inhibit E2F3 expression (53). Another study reported that reduction of PUM1 by RNA interference stabilized several mRNAs (33). PUMs were reported to repress the mRNA encoding CDKN1B tumor suppressor (32) and, unique to this mRNA, PUM1 was postulated to license microRNA mediated repression by disrupting basepairing between specific PUM and microRNA binding sites (32).

The role of deadenylases in yeast PUF repression suggested that human deadenylases might serve as PUM co-repressors. Humans possess multiple orthologs of the Pop2 and Ccr4 deadenylase enzymes (47). The human CNOT7 and CNOT8 proteins are related to yeast Pop2 while human CNOT6 and CNOT6L are orthologous to yeast Ccr4 (47,54-56). All four proteins have been reported to possess deadenylase activity (47,57-59). Like their yeast counterparts, CNOT7 and 8 form heterodimers with either CNOT6 or 6L, and these pairs assemble with human orthologs of the yeast Not proteins to form large multisubunit complexes referred to as Ccr4-Not (CNOT) complexes (60-62).

In this report, we explore the activities of human PUM1 and PUM2. We show that both PUMs are potent repressors that inhibit protein expression and reduce mRNA levels. We then investigate the mechanism of repression and show that purified PUM complexes contain CNOT deadenylases. Two deadenylase subunits interact directly with the PUMs. *In vivo*, we find that deadenylases are important PUM co-repressors and the poly(A) tail is necessary for efficient repression. We also present evidence for a poly(A) independent mechanism of PUM repression. This research reveals two modes of PUM repression and thereby enhances our understanding of their regulatory functions to control important biological processes.

Experimental Procedures

Plasmids. *Renilla* luciferase reporters (RnLUC) are based on psiCheck1 (Promega) with either three wild-type PRE or mutant PRE elements inserted into the Xho1 and Not1 sites in the 3'UTR. The PRE sequence is as follows, with the PRE underlined: 5'-TTGTTGTCGAAAATTGTACATAAGCCAA and the PREmt sequence is: 5'-TTGTTGTCGAAAATACAACATAAGCCAA. The altered specificity reporter, RnLUC 3xPRE UGG, was constructed with the following sequence: 5'-TTGTTGTCGAAAATTGGACATAAGCCAA. RnLUC HSL was created by replacing the cleavage/polyadenylation site from the psiCheck1 3' UTR with a histone stem loop (HSL) sequence from the human H1F3 gene. Two or four PRE sequences were inserted upstream of the HSL to create RnLUC 2xPRE HSL and RnLUC 4xPRE HSL. The Firefly luciferase (FfLUC) plasmid pGL4.13 (Promega) was used as a control.

To express proteins as Halotag fusions in human cells, CNOT6, CNOT6L, CNOT7, or CNOT8 were cloned into the vector pFN21A (Promega). Active site mutants, CNOT7 D40A E42A and CNOT8 D40A E42A were created by Quikchange mutagenesis (Stratagene). Full length human PUM1 or PUM2 open reading frames were cloned into pFN21A and site directed mutagenesis was used to create PUM1 R6as (N1043S Q1047E) and PUM2 R6as (N921S Q925E). Recombinant proteins were expressed as Halotag (HT) fusions from the vector pFN18A (Promega), including HT-CNOT6, HT-PUM1 RBD (aa828-1176), and HT-PUM2 RBD (aa705-1050). *Renilla* and firefly luciferase reporters for *Drosophila* cells were previously described (63). *Drosophila* protein expression constructs were made by inserting PUM1 coding sequence into pIZ V5 His6 (Invitrogen).

Cell culture and transfections. Human HEK293 cells were cultured at 37°C under 5% CO_2 in DMEM with glucose and 1 x Penicillin/Streptomycin/Glutamine and 10% FBS (Gibco). *Drosophila* D.mel-2 cells (Invitrogen) were cultured as previously described (63). Transfections of human cells were carried out with FuGENE HD (Promega) at 3:1 volume lipid: µg DNA. For luciferase assays, 2 x 10^4 cells were plated in white-walled 96-well plates and, after 24 hrs, were transfected with 100 ng/well of plasmid DNA. For RNA purifications and coimmunoprecipitations, 6 x 10^5 cells were transfected with 3 µg plasmid DNA 24 hours after seeding. D.mel-2 cells were transfected with Effectene (Qiagen) as previously described (63). For human PUM1 expression and repression assays, 400 ng PUM1 expression vector was included in the transfection with reporters.

Luciferase assays. Renilla (75 ng) and Firefly (25 ng) reporters were transfected into HEK293 cells. Forty-eight hours later, luciferase activity was measured with Dual-Glo reagent using a Glomax Multi+ luminometer (Promega). Relative light units (RLU) values were used to calculate a relative response ratio (RRR) by dividing the *Renilla* value from each well by the corresponding Firefly value. Percent repression was then calculated as: (%) = 100 x (1- $RRR_{variable}$/ $RRR_{control}$) where $RRR_{control}$ equals RRR RnLUC 3xPRE mt or RnLUC. A minimum of three replicates were used to calculate mean values and standard error of the mean. All results were verified in multiple independent experiments. Dual luciferase assays from *Drosophila* cells were performed as previously described (63).

RNA interference. PUMs were knocked down in HEK293 cells using On-target Plus Smartpool siRNAs for PUM1 (L-014179-00), PUM2 (L-014031-02), GAPDH, or non-targeting control siRNAs (Dharmacon). HEK293 cells (2 x 10^4 cells per well) were plated into a 96-well plate. After 24 hours, cells were transfected with 10 fmoles of siRNAs using Dharmafect-1

(Dharmacon). After 48 hours, reporters were transfected using FuGENE HD. Twenty-four hours later, cells were analyzed by Dual-Glo assay or whole cell lysates were prepared for western-blot analysis in TNEM (50 mM Tris-HCl pH 8.0, 0.5% (v/v) Nonidet-P40, 0.5 mM EDTA, 2 mM $MgCl_2$,) with 150 mM NaCl and protease inhibitors (1 mM PMSF, 50 μg/mL aprotinin, 50 μg/mL pepstatin, 50 μg/mL leupeptin).

RNAi in D.mel-2 cells was performed as previously described (63) using dsRNAs transcribed from PCR templates generated with the following oligonucleotides: LacZ control, forward primer, 5'-GGATCCTAATACGACTCACTATAGGGTGACGTCTCGTTGCTGCATAAAC, and reverse primer, 5'-GGATCCTAATACGACTCACTATAGGGGGCGTTAAAGTTGTTCTGCTTCATC, Pop2, forward primer, 5' -GGATCCTAATACGACTCACTATAGGGGGACACCGAGTTTCCAGGCG, reverse primer, 5'-GGATCCTAATACGACTCACTATAGGGGAAGAAGGCCATGCCCGTCAGC, Ccr4, forward primer, 5'-GGATCCTAATACGACTCACTATAGGGGGAAGTACGTCGATGGCTGTGC, reverse primer, 5'-GGATCCTAATACGACTCACTATAGGGCGAACGTATAGTTGGTGTGCGGCATT. The T7 promoter sequence is underlined and the gene specific region is in bold.

RNAi of POP2 and CCR4 was confirmed by measuring depletion of Halotag-deadenylase fusions. D.mel-2 cells were transfected with 100 ng pIZ HT-Pop2 or pIZ HT-CCR4 with 100 ng of control pIZ HT. 1 mL cell suspension was harvested and lysed for 1 hour on ice in TNEM with 150 mM NaCl. HT was labeled with fluorescent Halotag ligand, TMR (Promega) for 30 minutes. Lysates were analyzed by SDS-PAGE and fluorescence detection with a Typhoon Trio

Fluorescence Imager (GE). Depletion was calculated relative to samples treated with LacZ control dsRNA and normalized to HT internal control.

Coimmunoprecipitations. Plasmids expressing FLAG-tagged human PUM1 and PUM2 were transfected into 6 x 10^5 HEK293 cells with HT fusions of CNOT6, CNOT6L, CNOT7, or CNOT8. Cells were lysed in TNEM with 150 mM NaCl and protease inhibitors. HT fusions were labeled with TMR ligand and treated with 10 units RNase ONE and 4 μg RNase A (Promega). Extracts were then bound overnight with end-over-end rotation at 4°C to pre-equilibrated anti-FLAG beads (Sigma). Beads were washed twice with TNEM with 250 mM NaCl and once with TNEM with 500 mM NaCl. Bound protein complexes were eluted with FLAG peptide (Sigma) at 4°C and passed over Micro Bio-Spin columns (Bio-Rad) to collect eluates. Eluates were analyzed by SDS-PAGE and fluorescence emission at 580 nm on a Typhoon Trio to detect TMR-labeled Halotag fusions. Western blots were performed and probed with a monoclonal anti-FLAG M2 antibody (Sigma).

Western blotting. HEK293 cells were lysed in TNEM with 150mM NaCl and protease inhibitors. D.mel-2 lysates were prepared from 1 mL cell suspension in 75 μL TNEM with 150 mM NaCl and protease inhibitors, lysed on ice for 1 hour, and centrifuged to remove cell debris. Lysates were analyzed by western blotting with anti-V5. PUM1 antibodies were from Bethyl Laboratories (A300 201A) and Abcam (80216). PUM2 (K-14) antibody was from Santa Cruz Biotechnology (sc-31535) or Bethyl Laboratories (A300-202A). Antibody to GAPDH was obtained from Applied Biosystems (AM4300). T7 tag antibody was from Novagen (69522-3). V5 antibody was from Invitrogen (37-7500). HRP conjugated secondary antibodies were obtained from Thermo Scientific (anti-Mouse IgG, 31430) and KPL (Anti-Goat IgG, 14-13-06 and Anti-Rabbit IgG, 074-1516).

Purification of recombinant proteins. To purify PUM1 (aa828-1176) and PUM2 (aa705-1050) RNA binding domains and control CNOT6, pFN18A-based plasmids (Promega) encoding Halotag fusions of each protein were introduced into KRX *e. coli* strain (Promega) and induced with 0.1% (w/v) rhamnose for 12 hours at 20°C. Proteins were purified using Halolink resin (Promega). Beads were washed extensively with TNEM with 1000 mM NaCl and then equilibrated in TNEM with 250 mM NaCl. To confirm purification of the respective proteins, AcTEV protease (Invitrogen) was used to cleave CNOT6, PUM1 RBD and PUM2 RBD from an aliquot of the Halolink beads. The eluted proteins were analyzed by coomassie stained SDS-PAGE (Fig. 4B). The remaining Halolink bound proteins were used for Halotag pulldown assays. pMAL plasmids (NEB) encoding MBP tagged CNOT6, CNOT7, and CNOT8 were transformed into the BL21 Gold *E. coli* strain and induced with 0.3 mM IPTG for 16 hrs at 20°C. Proteins were purified with the Amylose affinity resin (NEB). Beads were washed three times with TNEM with 1000 mM NaCl and 1 mM DTT and three times with deadenylation buffer (50 mM Tris pH 8.0, 1 mM $MgCl_2$, 50 mM NaCl, 20% glycerol and 1 mM DTT). Proteins were eluted with 10 mM maltose in deadenylation buffer.

In vitro deadenylation assays. Deadenylase activity of purified wild type or mutant CNOT7 and CNOT8 enzymes was confirmed by incubating 1 μM of each enzyme with 200 fmol of a 36 nucleotide RNA substrate with a 5' Cy5 fluorescent label and, on the 3' end, a 10 nucleotide poly(A) tail (see PRE RNA sequence below) in 20 μL of deadenylation buffer (64). Control reactions contained 10 mM EDTA to chelate Mg^{2+}. Reactions were incubated at 30ºC for up to 120 minutes. Equal volume of 98% formamide and 20 mM EDTA was added, the samples were heated to 95ºC for 5 minutes and then resolved on a 10% polyacrylamide, 7 M urea gel. Products were detected using a Typhoon fluorescence imager.

In vitro binding of PUMs and CCR4-NOT deadenylase subunits. Recombinant prey proteins included MBP-CNOT7, MBP-CNOT8 or control MBP. For Halotag pulldown assays, 50 nM of prey protein was added to 50 μL of TNEM with 250 mM NaCl and 10 μLs of Halolink beads bound with HT-CNOT6, HT-CNOT7 or HT-CNOT8 bait proteins (1 μg each). Halolink beads alone served as a negative control. Binding reactions were incubated with rotation for 2 hours at 4°C. Beads were washed 4 times with 1 mL TNEM containing 500 mM NaCl and 0.5% Tween-20. Beads were collected by centrifugation at 1000 x g for 5 minutes. Bound proteins were eluted in 20 μL SDS-PAGE loading dye by heating at 95°C for 5 minutes. Fifty percent of eluted proteins were then analyzed by SDS-PAGE and western blotting using anti-MBP monoclonal antibody conjugated to horseradish peroxidase (NEB).

Gel shift assays. PRE RNA ligand, 5'-TTGTTGTCGAAAATTGTACATAAGCCAAAAAAAAAA, was labeled with Cy5 (Dharmacon). PRE mt RNA ligand, 5'-TTGTTGTCGAAAATACAACATAAGCCAAAAAAAAAA, was labeled with Dylight 650 (Dharmacon). RNA ligands were synthesized, deprotected and PAGE purified prior to gel shift assays. PUM1 RBD or PUM2 RBD were allowed to bind to 200 fmol (10 nM) of RNA ligand in deadenylation buffer for 30 minutes at 37°C. Samples were then analyzed on a 6% polyacrylamide gel with 1 x TB running buffer at 300 volts at 4°C. Gels were imaged with a Typhoon Trio.

Purification of PUMs and mass spectrometry. Halotag, HT-PUM1, or HT-PUM2, expressed from plasmid pFN21A, were purified using the Halotag Mammalian Pulldown system (Promega). T150 flasks were transfected with each plasmid and after 48 hours, cells were washed with phosphate buffered saline (PBS), and harvested at 2000 x g at 4°C. Cells were

suspended in 1 mL Mammalian Lysis Buffer with Protease Inhibitor Cocktail (Promega). Cells were passed through a 25 gauge needle 5 times, incubated for 5 minutes at 4°C, and then centrifuged for 5 minutes at 14000 rpm. Halolink beads were then diluted with TBS (100 mM Tris-HCl pH 7.5 and 150 mM NaCl) and incubated with the cell extract for 15 minutes at room temperature with rotation. Beads were washed three times with 10 ml TNEM with 250 mM NaCl, followed by three washes with the same buffer lacking IGEPAL. Proteins were eluted with 10 units of AcTEV protease (Invitrogen) in 20 mM Tris-HCl pH 8.0 and 300 mM NaCl. Peptides were prepared from each sample as follows: First, disulfide bonds were reduced with 2 mM DTT at 37°C for 30 minutes and blocked with 4 mM iodoacetamide at 23°C for 30 minutes in the dark. The blocking reaction was quenched by bringing the final concentration of DTT to 4 mM. Next, sequencing grade trypsin (Promega) at a 1:50 (mass:mass) enzyme to sample ratio was added and incubated overnight at 37°C. Peptides were then analyzed using nanoflow liquid chromatography (Waters) coupled to an ETD-enabled hybrid linear ion trap-orbitrap mass spectrometer (Thermo Scientific) via electrospray (65). Separation and data-dependent sampling conditions were used as previously described (66,67). Post-acquisition data processing was performed using DTA generator and the COMPASS software suite as previously described (68). Protein identifications were assigned by searching the human International Protein Index database with the peptide mass spectra from two independent analyses using the open mass spectrometry search algorithm (OMSSA) (67,69). A false discovery rate threshold of 1% was applied to filter false positive identifications (67,70). To eliminate contaminants that bind Halolink resin or Halotag, an identical analysis was performed on control Halotag purifications. All proteins detected in both the control and PUM complexes were excluded.

RNA purifications and cDNA preparation. RNA was purified from HEK293 cells harvested 48 hours after transfection using the Maxwell 16 simplyRNA LEV cells kit and a Maxwell 16 instrument (Promega). RNA was eluted in 50 μL nuclease free water and treated with Turbo DNase (Ambion).

For first strand cDNA synthesis, RNA (1000 ng) was annealed with random hexamers (500 ng) (IDT) at 70°C for 5 minutes and cooled on ice. Reverse transcription was performed in reaction buffer with 3 mM $MgCl_2$, 500 mM each dNTP, 0.5 μL RNasin Plus, and 1 μL GoScript reverse transcriptase (Promega). RT was omitted in control samples.

Quantitative PCR. Multiplexed quantitative PCR was used to detect *Renilla* and Firefly reporter mRNAs. Reactions were carried out in 25 μL reactions with the Plexor 2-step kit (Promega). 5 μL cDNA was combined with 2 x Plexor Master Mix (Promega) and 100 nM each of the fluorescent primers (Biosearch Technologies). Reactions were performed in triplicate using a CFX96 Real-Time PCR instrument (Bio-Rad). The conditions used were: 1) 95°C for 2 minutes, 2) 95°C for 5 seconds, 3) 60°C for 35 seconds. Steps 2-3 were repeated for a total of 40 cycles. Each reaction was subjected to thermal melting and curves gave single peaks with the expected melting temperature. Amplification efficiencies for each primer set were optimized at 100% efficiency. Cycle thresholds (Ct) were measured using CFX Manager software (Bio-Rad) and imported in Plexor Analysis Software (Promega). Data was analyzed by the comparative Ct method (71,72). Ct values were measured and normalized to an internal control firefly luciferase mRNA where $\Delta Ct = Ct_{\textit{Renilla}} - Ct_{\text{firefly}}$. Differences in mRNA levels were calculated using the ΔΔCt method whereby $\Delta\Delta Ct = \Delta Ct_{\text{target}} - \Delta Ct_{\text{control}}$. "Control" indicates RnLUC lacking PREs and "target" indicates RnLUC 3xPRE. Changes in mRNA expression are represented as fold

change values, where fold change = $2^{-\Delta\Delta Ct}$. From fold change we calculated percent repression, which equals 100 x (1- fold change).

To confirm RNAi depletion of deadenylase mRNAs, qPCR was carried out using GoTaq qPCR (Promega). Cycling conditions were as follows: (i) 95°C for 3 minutes, (ii) 95°C for 10 sec, (iii) 65°C for 30 sec, and (iv) 72°C for 40 sec. Steps (ii) through (iv) were repeated for a total of 40 cycles. Negative control reactions were performed in the absence of template or reverse transcriptase. Cycle thresholds (Ct) were measured using the CFX Manager software and analyzed using the ΔΔCt method (71,72). ΔCt was calculated by normalizing to the Rpl32 gene Ct values. We then calculated ΔΔCt as follows: ΔΔCt= ΔCt(target RNAi)- ΔCt (control RNAi). qPCR primer sequences are as follows: Firefly: Forward primer: 5'-dGATCCTCAACGTGCAAAAGAAGC-3', Reverse primer: 5'-d FAM-isoC-TCACGAAGGTGTACATGCTTTGG-3', Renilla: Forward primer: 5'-d CAL Fluor Orange 560-isoC-CGCAACTACAACGCCTACCTTC-3', Reverse primer: 5'-dCCCTCGACAATAGCGTTGGAAAA-3', Rpl32: Forward primer: 5'-dGCCCAAGGGTATCGACAACAG-3', Reverse primer: 5'-d GCACGTTGTGCACCAGGAAC-3', Dm Pop2: Forward primer: 5'-d TGGACAATGCCCTCGGCC-3', Reverse primer: 5'-d GGCCACATAGTGGTACTTCTGCACC-3', Dm Ccr4: Forward primer: 5'-d CTCGTCATACTCGGCCTCATGG-3', Reverse primer: 5'-d CGTAAAAATGCAGGCTGGTCG-3'.

Poly(A) selection and northern blot analysis. Total RNA samples from HEK293 cells expressing RnLUC, RnLUC 3xPRE, RnLUC 3xPREmt and FfLUC control were purified and then poly-adenylated mRNA was selected from 20 µg total RNA by the PolyAtract mRNA Isolation

System (Promega). RnLUC HSL and FfLUC RNAs were reverse transcribed and amplified by qPCR as described above.

For northern blots, mRNA samples were precipitated, suspended in 5 µL of 10 mM EDTA, 10 mM EGTA, and prepared for gel electrophoresis in 5% glycerol, 1 mM EDTA, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol FF with 1 x MOPS, 3.7% formaldehyde, and 25% formamide. Samples were boiled for 5 minutes and separated on a 1% agarose gel with formaldehyde in 1 x MOPS. Following transfer to NY+ membrane (Millipore), RNA was UV crosslinked and prehybridized in Ultra-Hyb buffer (Life Technologies) at 68°C for at least one hour. Northern probe templates for FfLUC and RnLUC were amplified by PCR using GoTaq (Promega) and the following primers*:*

FF fwd: 5'-CGTGGACGAGGTGCCTAAAG*,*

FF rev: 5' GGATCCTAATACGACTCACTATAGGTTACACGGCGATCTTGCCGC*,*

Rn fwd: 5'-CAAGCCCGACGTCGTCCAGATT*,*

Rn rev: 5'-GGATCCTAATACGACTCACTATAGGTTACTGCTCGTTCTTCAAGCACGC.

Reverse primers contained the T7 promoter sequence. Riboprobes were transcribed with UTP-[α-^{32}P] using T7 MaxiScript kit (Life Technologies) and purified by G25 Sephadex columns. Blots were hybridized with probes overnight at 68°C with rotation. Blots were washed twice for 15 minutes with 2X SSC (300 mM NaCl, 30 mM sodium citrate) + 0.1% SDS and twice for 30 minutes with 0.1% SSC (15 mM NaCl, 1.5 mM sodium citrate) + 0.1% SDS. Blots were exposed to a phosphorimager screen and visualized with a Typhoon Trio.

RESULTS

Human PUM1 and PUM2 reduce protein expression and mRNA levels. To study regulation by PUMs, we developed a luciferase reporter assay that recapitulates sequence specific repression. Three binding sites for PUM1 and PUM2, designated PUM Response Elements (PRE), were inserted into a minimal 3'UTR of an mRNA encoding *Renilla* luciferase (RnLUC 3xPRE, Fig. 1A). This PRE sequence UGUACAUA is a high affinity binding site for PUM1 and PUM2 (7). As a control for specificity, the UGU sequence of the PRE, which is crucial for PUM binding, was mutated to ACA to disrupt PUM binding (Fig. 1A, RnLUC 3xPREmt). Electrophoretic mobility shift assays confirm that PUM1 and PUM2 bind to the PRE with nearly equivalent affinity (Fig. 1B). Importantly, neither PUM bound the PREmt (Fig. 1B). As an additional control, a *Renilla* luciferase reporter lacking PRE sequences was tested (Fig. 1A, RnLUC).

Each reporter was transfected into the human HEK293 cell line. As an internal control, a plasmid encoding firefly luciferase was cotransfected (Fig. 1A, FfLUC). Expression of each luciferase was subsequently measured (Fig. 1C and D). *Renilla* expression from RnLUC 3xPRE was substantially repressed relative to RnLUC 3xPREmt or RnLUC (Fig. 1C). To normalize variations in transfection efficiency, the *Renilla* activity for each sample was divided by corresponding Firefly luciferase activity (Fig. 1D). From these values, we calculated a percent repression value, as a measure of PUMs repressive activity (Fig. 1E). The presence of the PRE elements in RnLUC 3xPRE elicited 71% repression relative to control reporters (Fig. 1F), indicating potent, specific repression by endogenous PUM1 and/or PUM2.

Having established that PRE dependent repression reduces protein output, we wished to determine if the effect is manifested by changes in mRNA level; therefore, we purified RNA and performed multiplexed quantitative reverse transcription polymerase chain reaction (qRT-PCR)

to measure levels of reporter mRNAs (Fig. 1F). RnLUC Ct values were normalized to the internal control, FfLUC, to yield a ΔCt value (71,72). From ΔCt values we calculated fold change for each sample, relative to negative control RnLUC (71,72). The fold change of RnLUC 3xPRE mRNA was 0.22, indicating it was reduced by 78% relative to RnLUC mRNA (Fig. 1G, 3xPRE). Consistent with repression by PUMs, mutation of the PREs alleviated regulation (Fig. 1G, 3xPREmt). Northern blotting was then performed using purified mRNA to visualize reporter transcripts (poly(A) affinity purification was necessary for detection). Detection of FfLUC served as an internal control. Quantification of the data revealed that RnLUC 3xPRE mRNA was reduced 74% relative to RnLUC and RnLUC 3xPREmt (Fig. 1G), concordant with qRT-PCR results (Fig. 1F). Together, these findings demonstrated that PUM repression of the PRE bearing reporter substantially reduces protein and mRNA levels, and the reporters provide sensitive sensors for post-transcriptional repression by PUMs.

Both PUM1 and PUM2 are expressed in HEK293 cells (Fig. 2A). To demonstrate that the PRE-dependent repression is caused by endogenous PUM1 and PUM2, each protein was depleted by RNA interference (RNAi). Transfection of non-targeting control siRNAs had no effect on PUM expression (Fig. 2A, Control). Treatment with siRNAs corresponding to PUM1 or PUM2 efficiently depleted the respective proteins (Fig. 2A, PUM1, PUM2). Treatment of cells with both PUM1 and PUM2 siRNAs substantially depleted both PUM1 and PUM2 (Fig.2A, PUM1+PUM2).

We then measured the effect of depletion of PUM1, PUM2, or both on reporter expression. The control siRNAs had no effect on repression of RnLUC 3xPRE (Fig. 2B, 65% repression) relative to mock transfection without siRNA (Fig. 2B, None). Likewise, transfection of siRNAs to GAPDH had no effect on repression (Fig. 2B, GAPDH). Depletion of each PUM

individually caused modest loss of repression (Fig. 2B). Depletion of both PUM1 and PUM2 together substantially reduced PUM repression to only 15% (Fig. 2B, PUM1+PUM2). We conclude that both PUMs repress the PRE-bearing reporter. We also tested the impact of over-expression of PUMs but did not observe enhancement of repression (data not shown), indicating that PUM expression is not limiting. Together, these observations indicate that both PUM1 and PUM2 cause PRE-dependent repression, and that they have overlapping regulatory roles. The results in Figures 1 and 2 validate the specificity and sensitivity of the PUM repression assay.

PUM1 and PUM2 repress individually. Having shown that PUMs have overlapping capabilities to repress, we next assessed whether PUM1 and PUM2 individually exhibit repressive activity. To do so, we created a new reporter that responds to exogenously introduced PUM1 or PUM2. First, each PUM was programmed to bind a new PRE sequence (designated PRE UGG) by altering the RNA recognition amino acids of the sixth PUF repeat (R6as)(Fig. 3A)(7,63). Importantly, wild type PUMs do not bind UGG efficiently (7,11). A corresponding reporter, RnLUC 3xPRE UGG, was created by changing the nucleobase at position 3 of the PRE from uracil to guanine (Fig. 3A and B). The reporters were then transfected into cells and regulation by endogenous PUMs or by PUM1 with altered specificity (PUM1 R6as) was measured. PUM1 R6as, fused to Halotag, was expressed from a transfected plasmid. As a control, a plasmid expressing only Halotag protein was introduced. As observed in Fig. 1, endogenous PUMs repressed the RnLUC 3xPRE but, importantly, did not affect RnLUC 3xPRE UGG, nor the negative controls RnLUC or RnLUC 3xPREmt (Fig. 3C, Halotag). Expression of PUM1 R6as specifically repressed the RnLUC 3xPRE UGG reporter by 64% (Fig. 3C). PUM1 R6as did not change repression of RnLUC 3xPRE by endogenous PUMs, nor did it regulate RnLUC or

RnLUC 3xPREmt (Fig. 3C). Next, we compared the repressive activity of PUM1 or PUM2 using the RnLUC 3xPRE UGG. PUM1 R6as repressed the reporter by 75% and PUM2 R6as repressed by 69% (Fig. 3D), relative to the Halotag control. We conclude that PUM1 and PUM2 can independently repress mRNAs, and that PUMs can be programmed to specifically repress new target mRNAs.

PUM1 and PUM2 interact with CCR4-NOT deadenylase complex subunits. We hypothesized that PUM1 and PUM2 may recruit co-repressor proteins to mediate repression. PUM complexes had not been previously biochemically analyzed. To identify co-repressors, we purified PUM1 and PUM2 complexes and identified associated proteins. First, PUMs were expressed in HEK293 cells as fusions to Halotag and affinity purified. Purified complexes were eluted and tryptic digests were then analyzed by nanoflow-reversed-phase liquid chromatography and electrospray ionization using a hybrid linear ion trap-orbitrap mass spectrometer. Peptide sequences and protein identifications were assigned by use of high accuracy mass spectral data (<10 ppm mass measurement) with a 1% false discovery rate cut off (67,70). To eliminate false-positives, an identical analysis was performed on control Halotag purifications; proteins detected in both the control and PUM complexes were excluded as contaminants. As a result of this analysis, multiple subunits of the CCR4-NOT (CNOT) deadenylase complex (61,73) were detected in purified PUM complexes including CNOT1, CNOT2, CNOT4 and CNOT10 (data not shown).

Association of CNOT subunits with PUMs prompted us to investigate interaction of deadenylase enzyme subunits with PUMs. The CNOT complex interacts with heterodimers formed by pairing of CNOT6 or CNOT6L with CNOT7 or CNOT8 deadenylases (47,60-62). To

analyze association of PUMs with these enzymes, FLAG-tagged PUM1 and PUM2 were expressed in cells that co-expressed Halotag fusion proteins of CNOT7, CNOT8, CNOT6 or CNOT6L. Cell extracts were prepared and treated with RNase One and RNase A to destroy RNA. Halotag fusions were fluorescently labeled with TMR fluor and detected in the cell lysates (Fig. 4A, Input). Next, PUM complexes were immunopurified using anti-FLAG monoclonal antibody and specifically eluted with FLAG peptide. Purification of PUM1 and PUM2 was confirmed by western blot of the eluates (Fig. 4A). CNOT6, CNOT6L and CNOT8 were strongly detected in both PUM1 and PUM2 eluates while CNOT7 was weakly detected (Fig. 4A). The interactions were specific, because none of the deadenylases or the Halotag control protein associated with the anti-FLAG resin (Fig. 4A, Control). This data demonstrates that PUMs associate with CNOT deadenylase complexes. Because the PUM-deadenylase association was detected in RNase treated extracts, protein interactions likely mediate the contacts and not RNA.

PUM1 and PUM2 bind the CNOT7 and CNOT8 deadenylases in vitro. To further investigate the interaction of PUMs with deadenylase enzymes, we performed *in vitro* protein interaction assays. As bait proteins, recombinant Halotag fusions of PUM1 and PUM2 were purified and immobilized to halolink beads (Fig. 4B). These proteins were active in RNA binding assays (Fig. 1B). As a positive control, a Halotag fusion of CNOT6 was also purified. Recombinant CNOT7 and CNOT8, fused to maltose binding protein (MBP) were then purified and used as prey proteins. First, the enzymatic activity of each deadenylase was demonstrated by deadenylating a 5' Cy5 fluorescently labeled RNA substrate with 10 nucleotide poly(A) tail (Fig. 4C). CNOT7 and CNOT8 progressively deadenylated the substrate over time. As a control, chelation of Mg^{2+}

with EDTA inactivated CNOT7 and CNOT8 (Fig. 4C, EDTA). Further, mutation of the magnesium coordinating residues (D40 and D42) within the active site of each enzyme to alanine blocked deadenylation (Fig. 4C).

Having demonstrated that CNOT7 and CNOT8 were active, we then measured binding to PUMs. Each prey was added to beads bound with CNOT6, PUM1, PUM2 or negative control beads. None of the prey proteins bound to control beads (Fig. 4D). The positive control, CNOT6, bound both CNOT7 and CNOT8, as expected (47), but not MBP (Fig. 4D). PUM1 and PUM2 bound to both CNOT7 and CNOT8, but not the control MBP (Fig. 4D). Therefore, human PUMs specifically interact with POP2 orthologs *in vitro*. Together with the results from co-immunoprecipitation studies (Fig. 4A), we conclude that PUMs bind either CNOT7 or CNOT8. We speculate that the preference for CNOT8 observed in Fig. 4A could result from additional factors *in vivo* that might modulate the interaction or differences in relative affinity. Because CNOT6/6L bind CNOT7 or CNOT8, their co-purification with PUMs is likely the result of heterodimerization.

Deadenylation inhibitors alleviate PUM repression. The observation that PUMs bind deadenylases suggests that deadenylation may be required for PUM mediated repression. To address this hypothesis, we used the observation that mutations in the catalytic residues of deadenylases render them inactive (Fig. 4C), and when over-expressed in cells, these mutants block deadenylation in a dominant negative manner (74-77). Therefore, we expressed mutant CNOT8 (CNOT8 mt) in which magnesium ion coordinating residues D40 and E42 were changed to alanine. The impact of these mutant deadenylases on PUM repression of RnLUC 3xPRE reporter was then measured. CNOT8 mt expression plasmid was transfected over a range from 0

to 85 nanograms (Fig. 5A). The CNOT8 mt protein was fused to Halotag to facilitate detection (Fig. 5B). A reciprocal gradient of the plasmid expressing only Halotag was used to balance transfections and Halotag alone served as a negative control. When Halotag alone was expressed (Fig. 5A, 0 ng CNOT8 mt), PUMs repressed the RnLUC 3xPRE by 77% relative to RnLUC, consistent with earlier observation (Fig. 1). Transfection of 20, 50 and 85 ng of the CNOT8 mt plasmid reduced PUM repression in a dose dependent manner to 58, 51, and 40%, respectively (Fig. 5A). The effect of CNOT8 mt was specific to PUM repression; neither RnLUC nor RnLUC 3xPREmt reporter was affected (Fig. 5A). Dose dependent expression of HT and HT-CNOT8 mt in these samples was confirmed by fluorescence detection (Fig. 5B). We conclude that CNOT8 mt has a dominant negative effect that inhibits repression by PUMs.

We next tested the ability of a catalytically inactivated mutant CNOT7 to affect PUM repression by the same strategy. Transfection of 20, 50, and 85 ng of CNOT7 mt expression plasmid reduced PUM repression from 78% to 56, 50 and 48%, respectively (Fig. 5C and 5D). Again, the effect was specific to the 3xPRE bearing reporter; RnLUC and RnLUC 3xPREmt reporters were not significantly affected. Together these results demonstrate that dominant negative mutant deadenylases block PUM repression, indicating that deadenylation plays an important role in PUM repression.

Depletion of deadenylases inhibits PUM repression. To corroborate the results above, we attempted to measure the impact of depletion of human deadenylases on PUM repression. Though we tested multiple siRNAs for each deadenylase, we were unable to substantially deplete CNOT7/8 and CNOT6/6L. Instead, we employed *Drosophila* D.mel-2 cells, which offer three advantages: 1) RNA interference elicited by dsRNA is highly efficient in these cells, 2)

Drosophila possess one copy each of POP2 (i.e. CAF1) and CCR4 (i.e. TWIN), thus circumventing the potential redundancy of deadenylases in human cells (47,78,79), and 3) human PUMs actively repress in D. mel-2 cells (see below).

We first confirmed the efficacy of RNAi mediated knockdown of deadenylases. To measure depletion of each protein, Halotag fusions of POP2 or CCR4 were co-expressed with a Halotag internal control. Cells were then treated with dsRNAs corresponding to either POP2 or CCR4 and, after 48 hours, levels of the Halotag fusion proteins were measured. POP2 and CCR4 were depleted by 99 and 94%, respectively (Fig. 6A), demonstrating efficient RNAi knockdown. We then tested the ability of human PUM1 to repress RnLUC 3xPRE in D.mel-2 cells. PUM1 repressed reporter protein expression by 45% relative to the empty expression vector (Fig. 6B). Simultaneous depletion of CCR4 and POP2 reduced repression to 28% (Fig. 6B). This effect was reflected at the mRNA level: PUM1 reduced RnLUC 3xPRE mRNA by 44% (Fig. 6C) and depletion of CCR4 and POP2 alleviated PUM1 mediated reduction of the mRNA to 19% (Fig. 6C). POP2 and CCR4 mRNAs were depleted from these samples by 82% and 94%, respectively, as ascertained by qRT-PCR. Therefore, the POP2 and CCR4 deadenylases are necessary for efficient repression by PUM1.

We sought to determine if the reduction in PUM1 repression was due to depletion of CCR4, POP2 or both. RNAi knockdown of CCR4 did not reduce PUM1 repression (Fig. 6D, CCR4, 58% repression) whereas PUM1 repression was significantly abrogated by knockdown of POP2 (Fig. 6D, POP2, 39% repression). This result may reflect the fact that POP2 is the predominant deadenylase in *Drosophila* (78,79). This finding supports the conclusion that deadenylases are important for PUM repression, and that PUM1 can repress by recruiting the deadenylase complex via a conserved interaction with POP2 orthologs.

PUMs also repress by a poly(A) independent mechanism. We next asked: is the poly(A) tail, and therefore deadenylation, absolutely necessary for repression by PUMs? Replication dependent histone mRNAs lack a poly(A) tail; rather, their 3' ends are formed by cleavage after a histone stem loop (HSL) structure (80). Translation of histone mRNAs is promoted by the 5' cap and HSL, and degradation occurs via the 5' decapping pathway (80). Consequently, the HSL provides a means of examining PUM repression in the absence of a poly(A) tail. We removed the cleavage/polyadenylation elements from the *Renilla* luciferase reporter and, in its place, inserted sequences encoding the HSL to drive 3' end formation of the RnLUC HSL reporter (Fig. 7A). To verify that the RnLUC HSL lacked a poly(A) tail, this mRNA was expressed in cells. As a positive control, the poly-adenylated RnLUC reporter was separately expressed. As an internal control, both samples also expressed the poly-adenylated FfLUC mRNA. Total RNA was purified from each sample and then mRNAs were then purified using oligo(dT) magnetic beads to enrich poly(A) mRNA. Using qRT-PCR, each mRNA was detected in the poly(A) selected fraction and normalized to the total amount. The poly(A) selected RNA contained less than 6% of RnLUC HSL mRNA whereas 100% of the control RnLUC mRNA was poly(A) selected (Fig. 7B). As expected, the FfLUC internal control was highly enriched in the poly(A) fraction (80-100%). These results confirm that at least 94% of the RnLUC HSL mRNA is not poly-adenylated.

To measure PUM repression, PREs were inserted into the 3' UTR to create RnLUC 2xPRE HSL and 4xPRE HSL (Fig. 7A). The two and four PREs conferred 22% and 57% repression, respectively (Figure 7C). To determine if repression of HSL reporters by PUMs affected their mRNA level, we measured the levels of each mRNA by qRT-PCR. PUM

repression did not reduce the RnLUC 2xPRE HSL reporter mRNA and, in fact, the 4xPRE HSL mRNAs was more abundant than RnLUC HSL mRNA (Fig. 7D). This indicates that PUM repression of the HSL reporters may occur at the translational level, rather than by direct activation of mRNA degradation pathways. From these data we conclude that PUMs can repress mRNAs lacking poly(A) tails and, therefore, can also repress by a deadenylation independent mechanism.

Discussion

Our results demonstrate that both human PUM1 and PUM2 are potent repressors that reduce levels of target mRNAs and cause a corresponding decrease in protein expression (Fig. 1). Endogenous PUMs have overlapping function and act redundantly to repress protein expression (Fig. 2). We show that human PUM1 and PUM2 repress autonomously and can be programmed to regulate new mRNAs, which offers potential therapeutic value for developing designer PUMs to reduce expression of deleterious genes (Fig. 3)(81,82). Furthermore, our results identify two modes of repression: deadenylation mediated repression and a deadenylation independent mechanism.

Our data provide the first evidence that human PUMs use deadenylase enzymes as co-repressors. PUMs physically associate with CNOT deadenylase subunits, including the four known deadenylase enzymes, CNOT6, 6L, 7 and 8 (Fig. 4), mediated by direct binding to human Pop2 orthologs, CNOT7 and CNOT8 (Fig. 4). The association of CNOT6 and 6L with PUMs is likely bridged via CNOT7 and CNOT8. Thus, we propose that PUMs recruit multiple deadenylase complexes to efficiently repress target mRNAs. The regulatory role of deadenylases in PUM repression is supported by the ability of dominant negative CNOT7 and CNOT8 mutants to inhibit PUM repression (Fig. 5). Importantly, these dominant negative mutants were

previously shown to inhibit deadenylation when expressed *in vivo* (74-77,79). Further support is provided by data showing that depletion of deadenylase enzymes reduces the magnitude of PUM repression (Fig. 6). Therefore, we conclude that deadenylation is necessary to achieve robust repression.

Removal of the poly(A) tail through concerted action of PUMs and deadenylases is anticipated to reduce translation efficiency and, at the same time, initiate degradation of the mRNA by either 5' decapping mediated decay, 3' decay by the exosome, or both pathways (83). This model is supported by our observation that protein and mRNA levels are concomitantly reduced by PUM repression. In accordance with this model, a previous study concluded that PUM1 promoted degradation of target mRNAs (33). It is noteworthy that we were unable to detect partially or fully deadenylated mRNAs, likely because these intermediates are unstable and low abundance.

Our results, combined with past data, indicate that PUF repression via deadenylation is a conserved mechanism. Correlations between deadenylation and PUF regulation were documented by studies in model organisms (29,35,38-41). Compelling evidence came from yeast, where PUFs were shown to accelerate mRNA deadenylation and degradation (42-45,84,85). The Ccr4-Pop2 deadenylase is required for repression by yeast PUFs (43-45). In all cases, including human PUMs (Fig. 4), the highly conserved PUF RNA binding domain was sufficient for interaction with Pop2 orthologs (35,43,45,86); therefore, the PUF repeat domain is likely responsible for deadenylation mediated repression.

While deadenylases are important for PUM repression, several observations provide evidence for a second poly(A) independent repression mechanism. First, dominant negative CNOT7/8 mutants do not completely block repression *in vivo* (Fig. 5). Second, depletion of

deadenylases does not fully alleviate Pum repression (Fig. 6). The third, more telling finding is that PUMs repress target mRNAs with a 3' HSL, indicating that a poly(A) tail - and consequently deadenylation - is not absolutely essential. Taken together, these data support an additional deadenylation independent repression mechanism.

Deadenylation dependent and independent mechanisms may function together to achieve maximal regulation. Indeed, the magnitude of repression of HSL mRNAs was less than that observed with polyadenylated reporter. Our results are reminiscent of a study that analyzed repression by the *Drosophila* PUF protein, Pumilio, wherein embryos were injected with reporter mRNAs either bearing a poly(A) tail or lacking a tail. Pumilio repressed the poly(A) mRNA most efficiently and, to a lesser degree, the tail-less RNA (40). Other mRNA regulators have also been reported to repress by deadenylation independent mechanisms. For instance, artificial tethering of the miRNA effector protein GW182 or the CNOT complex can inhibit HSL reporters, suggesting that PUM recruitment of CNOT might cause translation repression independent of deadenylation (87,88)

How do PUMs cause deadenylation independent PUM repression? In addition to deadenylation, PUMs could activate another mRNA decay step, such as decapping; though the observation that the PRE containing HSL target mRNA were not degraded argues against this hypothesis (Fig. 7). Alternatively, PUMs might interfere with translation, supported by work in model organisms indicating that PUFs can inhibit translation (48,51,89). Germane to this idea, PUFs were recently reported to bind to a translation elongation factor (51). Furthermore, we recently characterized conserved repression domains in the N-terminus of *Drosophila* and human PUFs that may elicit deadenylation independent repression (63). Future investigations will evaluate these possible mechanisms.

References

1. Moore, M. J. (2005) *Science* **309**, 1514-1518
2. Wickens, M., Bernstein, D. S., Kimble, J., and Parker, R. (2002) *Trends Genet* **18**, 150-157
3. Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997) *Rna* **3**, 1421-1433
4. Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickens, M. P. (1997) *Nature* **390**, 477-484
5. Wharton, R. P., Sonoda, J., Lee, T., Patterson, M., and Murata, Y. (1998) *Mol Cell* **1**, 863-872
6. Edwards, T. A., Pyle, S. E., Wharton, R. P., and Aggarwal, A. K. (2001) *Cell* **105**, 281-289
7. Wang, X., McLachlan, J., Zamore, P. D., and Hall, T. M. (2002) *Cell* **110**, 501-512
8. Wang, X., Zamore, P. D., and Hall, T. M. (2001) *Mol Cell* **7**, 855-865
9. Spassov, D. S., and Jurecic, R. (2002) *Gene* **299**, 195-204
10. Spassov, D. S., and Jurecic, R. (2003) *IUBMB Life* **55**, 359-366
11. Cheong, C. G., and Hall, T. M. (2006) *Proc Natl Acad Sci U S A* **103**, 13635-13639
12. Jenkins, H. T., Baker-Wilding, R., and Edwards, T. A. (2009) *J Struct Biol* **167**, 271-276
13. Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., and Gerber, A. P. (2008) *PLoS One* **3**, e3164
14. Spassov, D. S., and Jurecic, R. (2003) *Blood Cells Mol Dis* **30**, 55-69
15. Lin, H., and Spradling, A. C. (1997) *Development* **124**, 2463-2476
16. Parisi, M., and Lin, H. (1999) *Genetics* **153**, 235-250
17. Wharton, R. P., and Struhl, G. (1991) *Cell* **67**, 955-967
18. Asaoka-Taguchi, M., Yamada, M., Nakamura, A., Hanyu, K., and Kobayashi, S. (1999) *Nat Cell Biol* **1**, 431-437
19. Barker, D. D., Wang, C., Moore, J., Dickinson, L. K., and Lehmann, R. (1992) *Genes Dev* **6**, 2312-2326
20. Crittenden, S. L., Bernstein, D. S., Bachorik, J. L., Thompson, B. E., Gallegos, M., Petcherski, A. G., Moulder, G., Barstead, R., Wickens, M., and Kimble, J. (2002) *Nature* **417**, 660-663
21. Dubnau, J., Chiang, A. S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., Broger, C., and Tully, T. (2003) *Curr Biol* **13**, 286-296
22. Lehmann, R., and Nusslein-Volhard, C. (1987) *Nature* **329**, 167-170
23. Mee, C. J., Pym, E. C., Moffat, K. G., and Baines, R. A. (2004) *J Neurosci* **24**, 8695-8703
24. Menon, K. P., Sanyal, S., Habara, Y., Sanchez, R., Wharton, R. P., Ramaswami, M., and Zinn, K. (2004) *Neuron* **44**, 663-676
25. Nakahata, S., Kotani, T., Mita, K., Kawasaki, T., Katsu, Y., Nagahama, Y., and Yamashita, M. (2003) *Mech Dev* **120**, 865-880
26. Rouhana, L., Wang, L., Buter, N., Kwak, J. E., Schiltz, C. A., Gonzalez, T., Kelley, A. E., Landry, C. F., and Wickens, M. (2005) *Rna* **11**, 1117-1130

27. Schweers, B. A., Walters, K. J., and Stern, M. (2002) *Genetics* **161**, 1177-1185
28. Vessey, J. P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M. A., and Macchi, P. (2006) *J Neurosci* **26**, 6496-6508
29. Wreden, C., Verrotti, A. C., Schisa, J. A., Lieberfarb, M. E., and Strickland, S. (1997) *Development* **124**, 3015-3023
30. Chen, D., Zheng, W., Lin, A., Uyhazi, K., Zhao, H., and Lin, H. (2012) *Curr Biol* **22**, 420-425
31. Xu, E. Y., Chang, R., Salmon, N. A., and Reijo Pera, R. A. (2007) *Mol Reprod Dev* **74**, 912-921
32. Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A., Elkon, R., and Agami, R. (2010) *Nat Cell Biol* **12**, 1014-1020
33. Morris, A. R., Mukherjee, N., and Keene, J. D. (2008) *Mol Cell Biol* **28**, 4093-4103
34. Quenault, T., Lithgow, T., and Traven, A. (2011) *Trends Cell Biol* **21**, 104-112
35. Suh, N., Crittenden, S. L., Goldstrohm, A., Hook, B., Thompson, B., Wickens, M., and Kimble, J. (2009) *Genetics* **181**, 1249-1260
36. Kaye, J. A., Rose, N. C., Goldsworthy, B., Goga, A., and L'Etoile, N. D. (2009) *Neuron* **61**, 57-70
37. Pique, M., Lopez, J. M., Foissac, S., Guigo, R., and Mendez, R. (2008) *Cell* **132**, 434-448
38. Ahringer, J., and Kimble, J. (1991) *Nature* **349**, 346-348
39. Ahringer, J., Rosenquist, T. A., Lawson, D. N., and Kimble, J. (1992) *Embo J* **11**, 2303-2310
40. Chagnovich, D., and Lehmann, R. (2001) *Proc Natl Acad Sci U S A* **98**, 11359-11364
41. Gamberi, C., Peterson, D. S., He, L., and Gottlieb, E. (2002) *Development* **129**, 2699-2710
42. Olivas, W., and Parker, R. (2000) *Embo J* **19**, 6602-6611
43. Goldstrohm, A. C., Hook, B. A., Seay, D. J., and Wickens, M. (2006) *Nat Struct Mol Biol* **13**, 533-539
44. Goldstrohm, A. C., Seay, D. J., Hook, B. A., and Wickens, M. (2007) *J Biol Chem* **282**, 109-114
45. Hook, B. A., Goldstrohm, A. C., Seay, D. J., and Wickens, M. (2007) *J Biol Chem* **282**, 15430-15438
46. Blewett, N. H., and Goldstrohm, A. C. (2012) *Mol Cell Biol*
47. Goldstrohm, A. C., and Wickens, M. (2008) *Nat Rev Mol Cell Biol* **9**, 337-344
48. Chritton, J. J., and Wickens, M. (2010) *Rna* **16**, 1217-1225
49. Cao, Q., and Richter, J. D. (2002) *Embo J* **21**, 3852-3862
50. Cho, P. F., Gamberi, C., Cho-Park, Y. A., Cho-Park, I. B., Lasko, P., and Sonenberg, N. (2006) *Curr Biol* **16**, 2035-2041
51. Friend, K., Campbell, Z. T., Cooke, A., Kroll-Conner, P., Wickens, M. P., and Kimble, J. (2012) *Nat Struct Mol Biol* **19**, 176-183
52. Lee, M. H., Hook, B., Pan, G., Kershner, A. M., Merritt, C., Seydoux, G., Thomson, J. A., Wickens, M., and Kimble, J. (2007) *PLoS Genet* **3**, e233
53. Miles, W. O., Tschop, K., Herr, A., Ji, J. Y., and Dyson, N. J. (2012) *Genes Dev* **26**, 356-368

54. Albert, T. K., Lemaire, M., van Berkum, N. L., Gentz, R., Collart, M. A., and Timmers, H. T. (2000) *Nucleic Acids Res* **28**, 809-817
55. Dupressoir, A., Morel, A. P., Barbot, W., Loireau, M. P., Corbo, L., and Heidmann, T. (2001) *BMC Genomics* **2**, 9
56. Draper, M. P., Salvadore, C., and Denis, C. L. (1995) *Mol Cell Biol* **15**, 3487-3495
57. Morita, M., Suzuki, T., Nakamura, T., Yokoyama, K., Miyasaka, T., and Yamamoto, T. (2007) *Mol Cell Biol* **27**, 4980-4990
58. Bianchin, C., Mauxion, F., Sentis, S., Seraphin, B., and Corbo, L. (2005) *Rna* **11**, 487-494
59. Viswanathan, P., Ohn, T., Chiang, Y. C., Chen, J., and Denis, C. L. (2004) *J Biol Chem* **279**, 23988-23995
60. Collart, M. A., and Panasenko, O. O. (2012) *Gene* **492**, 42-53
61. Lau, N. C., Kolkman, A., van Schaik, F. M., Mulder, K. W., Pijnappel, W. W., Heck, A. J., and Timmers, H. T. (2009) *Biochem J* **422**, 443-453
62. Wagner, E., Clement, S. L., and Lykke-Andersen, J. (2007) *Mol Cell Biol* **27**, 1686-1695
63. Weidmann, C. A., and Goldstrohm, A. C. (2012) *Mol Cell Biol* **32**, 527-540
64. Goldstrohm, A. C., Hook, B. A., and Wickens, M. (2008) *Methods Enzymol* **448**, 77-106
65. McAlister, G. C., Berggren, W. T., Griep-Raming, J., Horning, S., Makarov, A., Phanstiel, D., Stafford, G., Swaney, D. L., Syka, J. E., Zabrouskov, V., and Coon, J. J. (2008) *J Proteome Res* **7**, 3127-3136
66. Lee, M. V., Topper, S. E., Hubler, S. L., Hose, J., Wenger, C. D., Coon, J. J., and Gasch, A. P. (2011) *Mol Syst Biol* **7**, 514
67. Phanstiel, D. H., Brumbaugh, J., Wenger, C. D., Tian, S., Probasco, M. D., Bailey, D. J., Swaney, D. L., Tervo, M. A., Bolin, J. M., Ruotti, V., Stewart, R., Thomson, J. A., and Coon, J. J. (2011) *Nat Methods* **8**, 821-827
68. Wenger, C. D., Phanstiel, D. H., Lee, M. V., Bailey, D. J., and Coon, J. J. (2011) *Proteomics*
69. Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. (2004) *J Proteome Res* **3**, 958-964
70. McAlister, G. C., Phanstiel, D., Wenger, C. D., Lee, M. V., and Coon, J. J. (2010) *Anal Chem* **82**, 316-322
71. Livak, K. J., and Schmittgen, T. D. (2001) *Methods* **25**, 402-408
72. Schmittgen, T. D., and Livak, K. J. (2008) *Nat Protoc* **3**, 1101-1108
73. Goldstrohm, A. C., and Wickens, M. (2008) *Nature Reviews Molecular and Cell Biology* **9**, 337-344
74. Piao, X., Zhang, X., Wu, L., and Belasco, J. G. (2010) *Mol Cell Biol* **30**, 1486-1494
75. Yamashita, A., Chang, T. C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C. Y., and Shyu, A. B. (2005) *Nat Struct Mol Biol* **12**, 1054-1063
76. Zheng, D., Ezzeddine, N., Chen, C. Y., Zhu, W., He, X., and Shyu, A. B. (2008) *J Cell Biol* **182**, 89-101
77. Chen, C. Y., Zheng, D., Xia, Z., and Shyu, A. B. (2009) *Nat Struct Mol Biol* **16**, 1160-1166

78. Temme, C., Zaessinger, S., Meyer, S., Simonelig, M., and Wahle, E. (2004) *Embo J* **23**, 2862-2871
79. Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., and Wahle, E. (2010) *Rna* **16**, 1356-1370
80. Marzluff, W. F., Wagner, E. J., and Duronio, R. J. (2008) *Nat Rev Genet* **9**, 843-854
81. Wang, Y., Cheong, C. G., Hall, T. M., and Wang, Z. (2009) *Nat Methods* **6**, 825-830
82. Dong, S., Wang, Y., Cassidy-Amstutz, C., Lu, G., Bigler, R., Jezyk, M. R., Li, C., Hall, T. M., and Wang, Z. (2011) *J Biol Chem*
83. Parker, R., and Song, H. (2004) *Nat Struct Mol Biol* **11**, 121-127
84. Houshmandi, S. S., and Olivas, W. M. (2005) *Rna* **11**, 1655-1666
85. Jackson, J. S., Jr., Houshmandi, S. S., Lopez Leban, F., and Olivas, W. M. (2004) *Rna* **10**, 1625-1636
86. Kadyrova, L. Y., Habara, Y., Lee, T. H., and Wharton, R. P. (2007) *Development* **134**, 1519-1527
87. Chekulaeva, M., Mathys, H., Zipprich, J. T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011) *Nat Struct Mol Biol* **18**, 1218-1226
88. Cooke, A., Prigge, A., and Wickens, M. (2010) *J Biol Chem* **285**, 28506-28513
89. Cao, Q., Padmanabhan, K., and Richter, J. D. (2010) *Rna* **16**, 221-227

Acknowledgments
We thank Dr. David Turner for comments on this manuscript, Nathan Blewett and Nathan Raynard for helpful suggestions, and May Tsoi for assisting with a multitude of lab tasks. This research was supported in part by a grant from the University of Michigan, Center for Genetics in Health and Medicine, Genomics and Genetics Grant and a grant from Promega. C.A.W. is supported by the Michigan Predoctoral Training Program in Genetics through NIH National Research Service Award 5T32GM007544-33. Conflict of interest statement: T.L.S, spouse of A.C.G., is an employee of Promega Corporation.

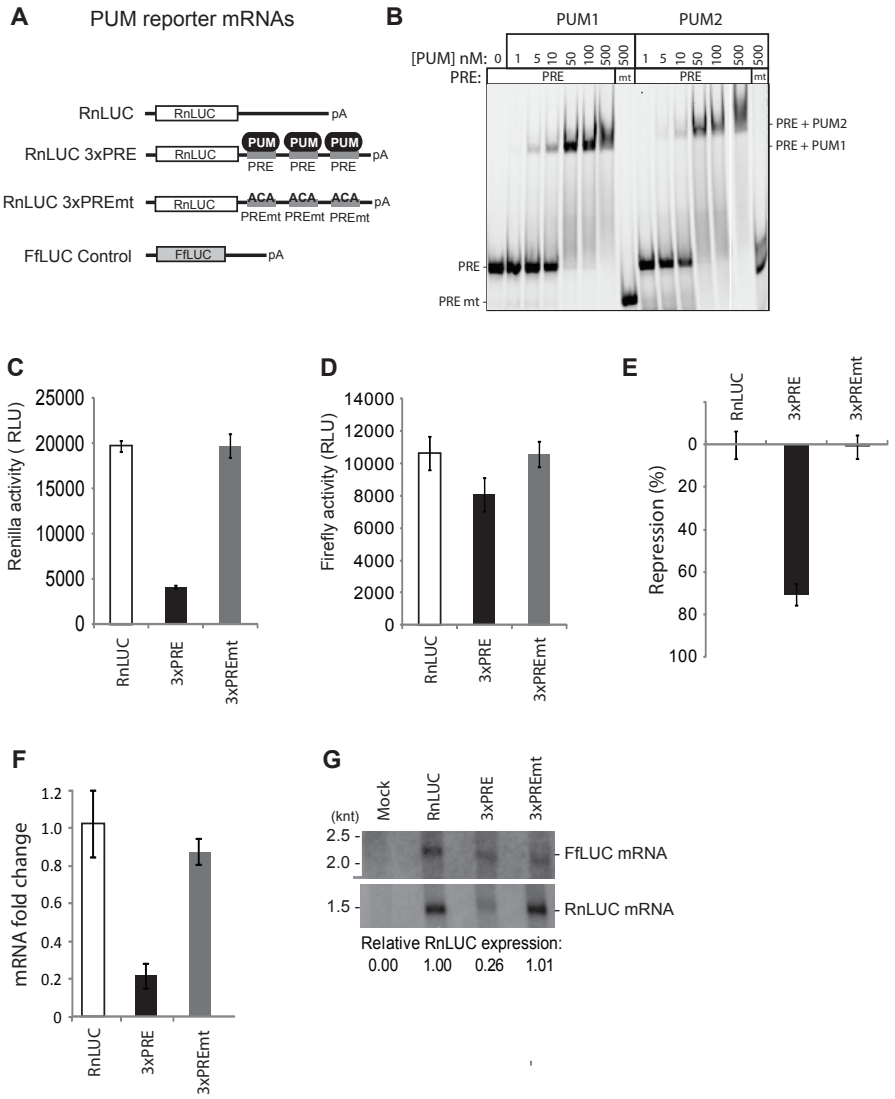


Figure 3.1. Repression by human PUM1 and PUM2 reduces protein and mRNA levels.

A. A luciferase reporter assay was developed to study PUM repression in human cells. Three Pumilio Response Elements (PRE) were inserted into the 3' UTR of *Renilla* luciferase (RnLUC) to create the reporter RnLUC 3xPRE. As a control, those sites were mutated in RnLUC 3xPREmt to block PUM binding. Firefly luciferase (FfLUC) was cotransfected as an internal control. *B.* Gel shift assay showing that PUM1 and PUM2 bind to PRE RNA with nearly equal affinity, but do not bind the mutant PRE (PREmt). PRE RNA was labeled at the 5' end with Cy5 fluor, whereas PREmt had a Dylight650 fluor, accounting for the difference in mobilities. *C.* Graph of *Renilla* reporter activity. Relative luminescence units (RLU) are normalized to the activity of the transfection efficiency control, FfLUC, shown in *D.* *E.* Percent repression of luciferase expression from RnLUC 3xPRE and RnLUC 3xPREmt, calculated relative to RnLUC. *F.* Reporter mRNA levels were measured by multiplexed qRT-PCR and fold changes, relative to RnLUC and normalized to FfLUC, are plotted. *G.* Northern blot detection of *Renilla* reporters and the control, FfLUC, from equal amounts of poly(A) selected mRNA. Relative expression levels were calculated by normalizing *Renilla* mRNA in each sample to the FfLUC mRNA in that sample. mRNA from mock transfected cells demonstrates specificity of probes. Mean values are graphed and standard error of the mean is indicated.

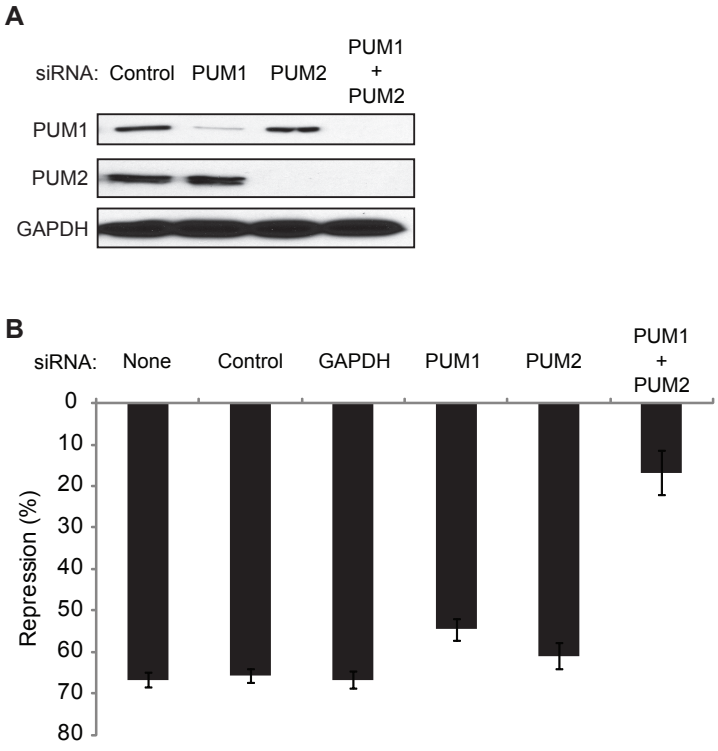


Figure 3.2. PUM1 and PUM2 repress RnLUC 3xPRE.

A. Endogenous PUM1 and PUM2 were depleted from HEK293 cells using siRNAs, as indicated at the top. Control indicates non-targeting control siRNAs. Western blot of PUMs using specific antibodies. GAPDH western blot on the same samples served as a loading control. *B.* Graph of percent repression of RnLUC 3xPRE, calculated relative to RnLUC, in samples treated with the siRNAs indicated at the top. "None" designates that the samples were mock transfected without siRNAs. "Control" indicates non-targeting control siRNAs. Mean values are graphed and standard error of the mean is indicated.

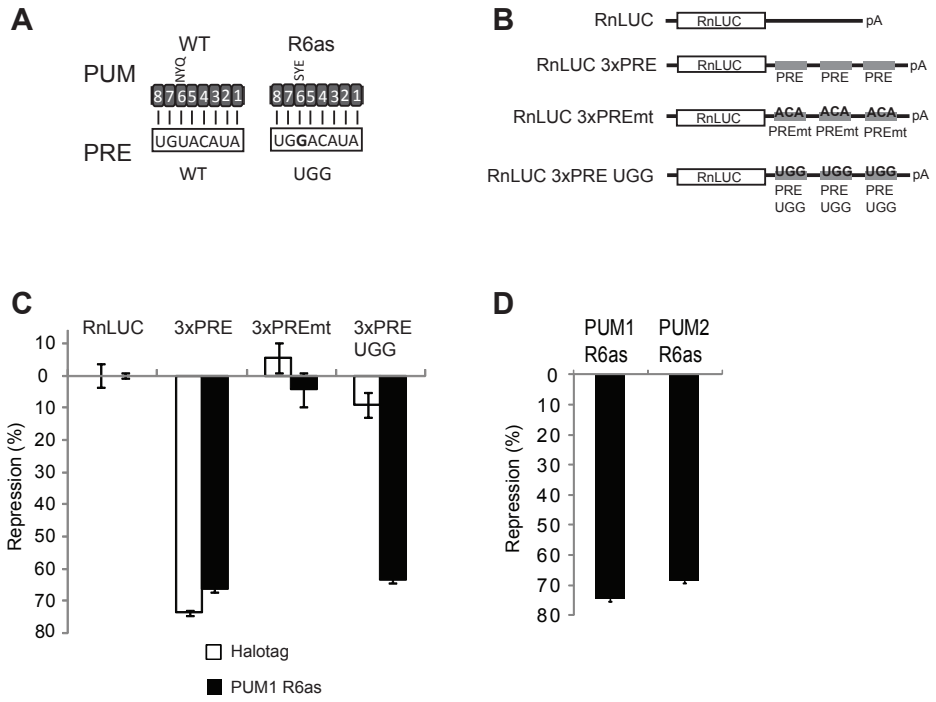


Figure 3.3. PUM1 and PUM2 repress individually.

A. Wild type PUMs bind to the wild type (WT) PRE sequence. Altered specificity PUMs (R6as) were created by changing RNA recognition amino acids of repeat 6 from NYQ to SYE, thereby binding to the altered PRE with a UGG sequence. Numbers indicate the PUF repeats, aligned to corresponding PRE ribonucleotide. *B.* Diagram of *Renilla* luciferase (RnLUC) reporters with three PREs (RnLUC 3xPRE), with mutant PREs (UGU changed to ACA) that cannot bind PUMs (RnLUC 3xPREmt) or reporter that is bound specifically by alter specificity PUMs (RnLUC 3xPRE UGG). *C.* Regulation of each reporter calculated as percent repression relative to RnLUC reporter. Endogenous PUM1 and PUM2 repress RnLUC 3xPRE but not RnLUC 3xPRE UGG (Halotag samples). Halotag was expressed as a negative control. PUM1 R6as, expressed as a Halotag fusion, specifically represses RnLUC 3xPRE UGG. *D.* PUM1 R6as and PUM2 R6as repressed the RnLUC 3xPRE UGG reporter. Both PUM1 R6as and PUM2 R6as proteins were expressed as fusions to Halotag. Percent repression was calculated relative to reporter expression in samples transfected with Halotag control. Mean values are graphed and standard error of the mean is indicated.

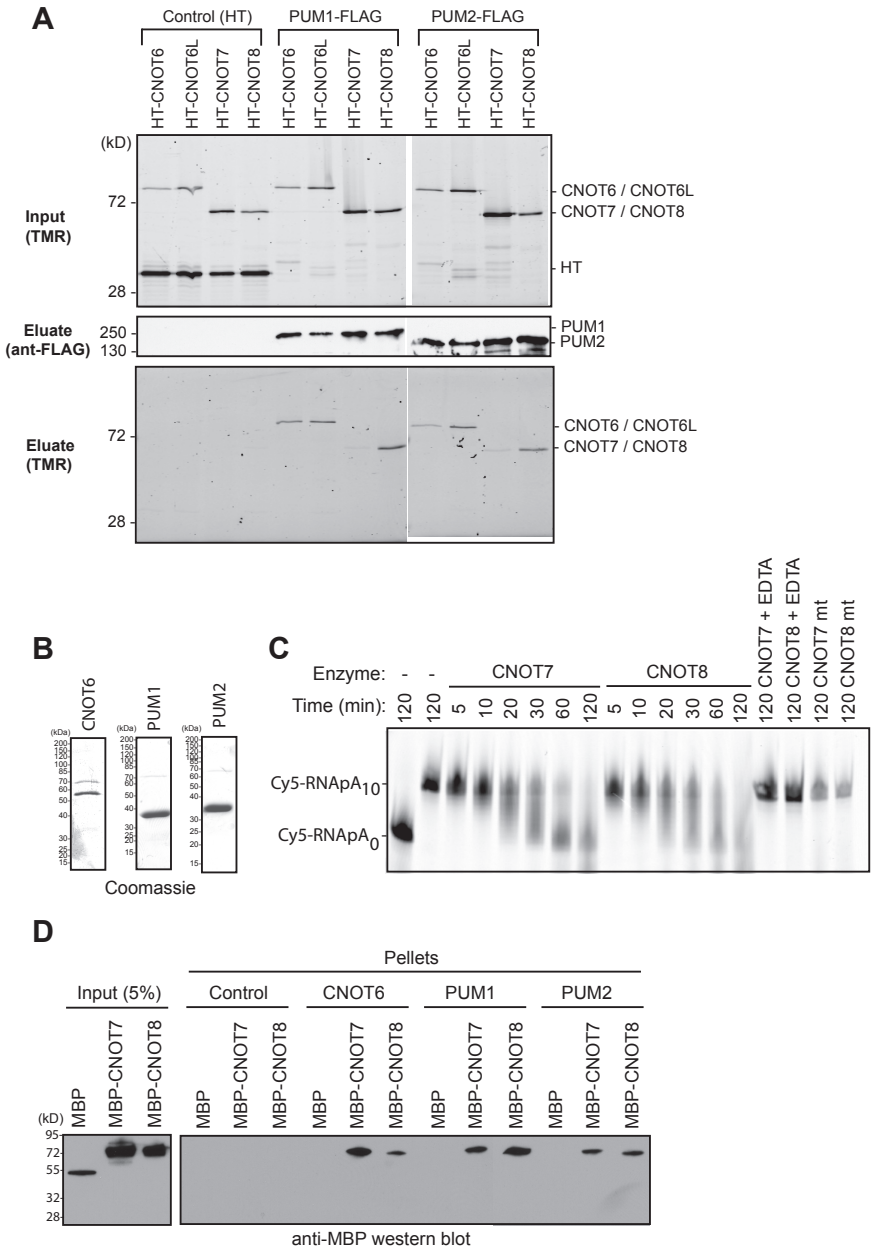


Figure 3.4. PUM1 and PUM2 interact with deadenylase subunits of the CNOT complex.

A. Deadenylases, fused to Halotag, coimmunoprecipitate (Eluate) with FLAG-tagged PUM1 and PUM2 from RNase treated extracts (Input). As a negative control (Control), mock immunoprecipitations were performed with anti-FLAG beads from samples expressing Halotag (HT) protein and Halotag deadenylase fusion proteins. Proteins were detected in input extracts or purified FLAG eluates by fluorescence labeling with the Halotag ligand TMR or by anti-FLAG western blot. *B.* Coomassie staining of recombinant, purified bait proteins: CNOT6, PUM1, and PUM2. PUMs were active for RNA binding (Fig. 1B). *C. In vitro* deadenylation assay using wild type CNOT7 and CNOT8 or mutant CNOT7 mt and CNOT8 mt with Cy5 labeled RNA substrate with 10 nucleotide poly(A) tail (Cy5-RNApA$_{10}$) or, as a marker, substrate lacking a tail (Cy5-RNApA$_0$). EDTA was added as a negative control to chelate Mg^{2+} and thus inhibit deadenylation. *D*. Western blot (anti-MBP) of *in vitro* binding of recombinant, purified PUM1 and PUM2 to CNOT7 and CNOT8. Halolink bound PUM1 and PUM2 were incubated with MBP fusions of CNOT7 or CNOT8. As a positive control, CNOT7 and CNOT8 interacted with CNOT6. Halolink beads alone (Control) and MBP served as negative controls.

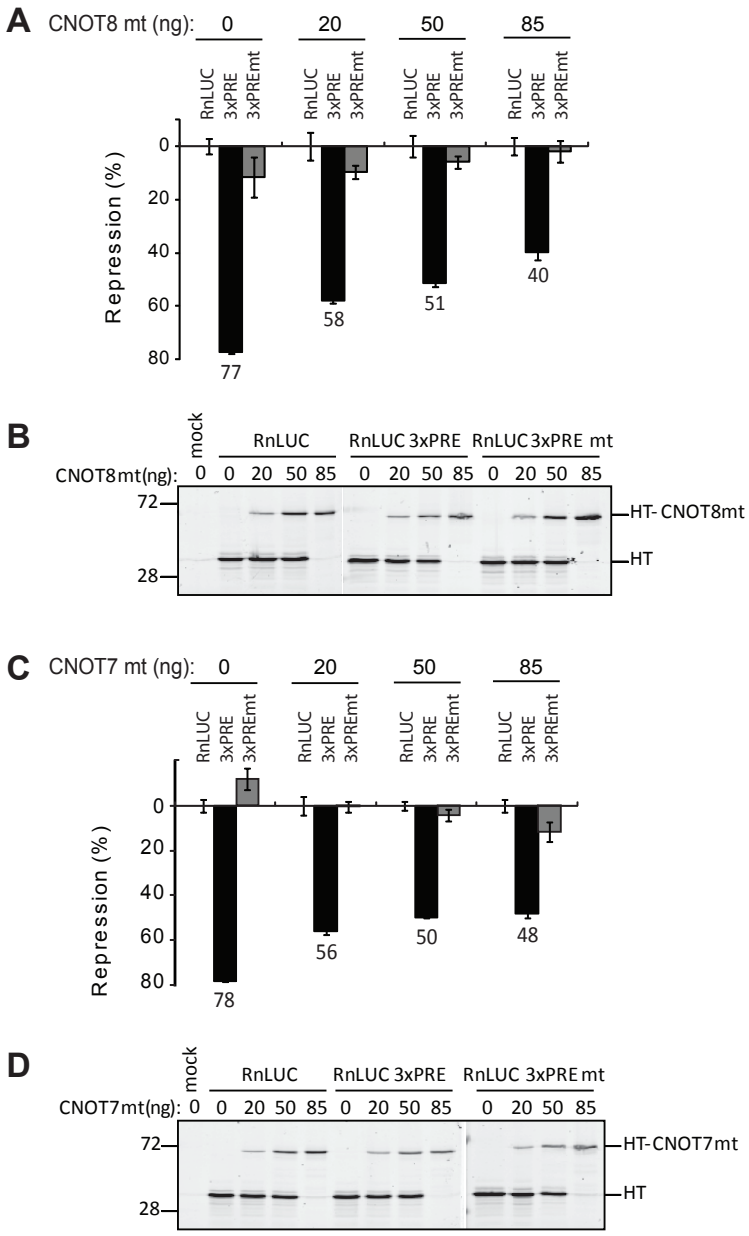


Figure 3.5. Dominant negative deadenylases alleviate PUM repression.

A. Expression of a dominant negative CNOT8 mt inhibits PUM repression in HEK293 cells. Graph of percent repression, relative to the RnLUC, of the indicated reporters in cells transfected with increasing amounts of plasmid expressing CNOT8 mt protein. *B*. Expression of the Halotag-CNOT8 mt fusion protein was confirmed by fluorescent TMR labeling and detection on SDS-PAGE gel from the samples in panel A. Halotag alone was used to balance transfected mass of DNA and is therefore also detected. *C*. Dominant negative CNOT7 mt inhibits PUM repression. Graph of percent repression relative to RnLUC control of the indicated reporters in cells transfected with increasing amounts of plasmid expressing CNOT7 mt protein. *D.* Expression of Halotag-CNOT7 mt fusions was confirmed by TMR detection on SDS-PAGE from the samples in panel C. Halotag alone was used to balance transfected mass of DNA and is therefore also detected. In all graphs, mean values are graphed with standard error of the mean.

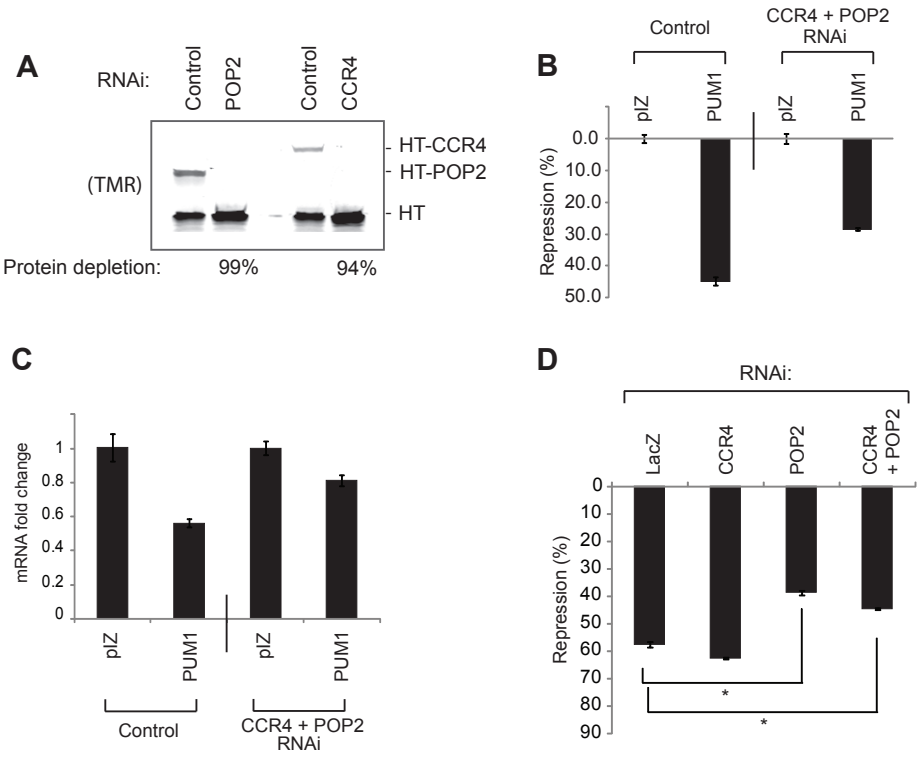


Figure 3.6. Depletion of deadenylases reduces PUM repression.

A. RNAi depletion of Halotag fusions of *Drosophila* CCR4 (HT-CCR4) and POP2 (HT-POP2) deadenylases in D.mel-2 cells, assayed by SDS-PAGE and TMR fluorescence detection. Halotag (HT) alone served as an internal control. Percent knockdown of each protein is indicated below the figure. *B*. Human PUM1 represses RnLUC 3xPRE reporter by 45% in D.mel-2 cells, relative to empty expression vector, pIZ. RNAi depletion of endogenous CCR4 and POP2 reduced repression to 28%. *C*. RnLUC 3xPRE mRNA levels were measured from samples in panel B using multiplexed qRT-PCR to determine the fold change in mRNA levels relative to empty expression vector, pIZ. PUM1 reduced mRNA levels 44% on the control sample versus only 19% when deadenylases were depleted by RNAi. *D*. RNAi depletion of endogenous POP2 inhibits PUM1 repression whereas knockdown of endogenous CCR4 does not. Non-targeting double-stranded RNA corresponding to the bacterial LacZ gene served as a negative control. Statistical significance is indicated with *, representing $P<0.0001$ by a two-tailed, unpaired t test.

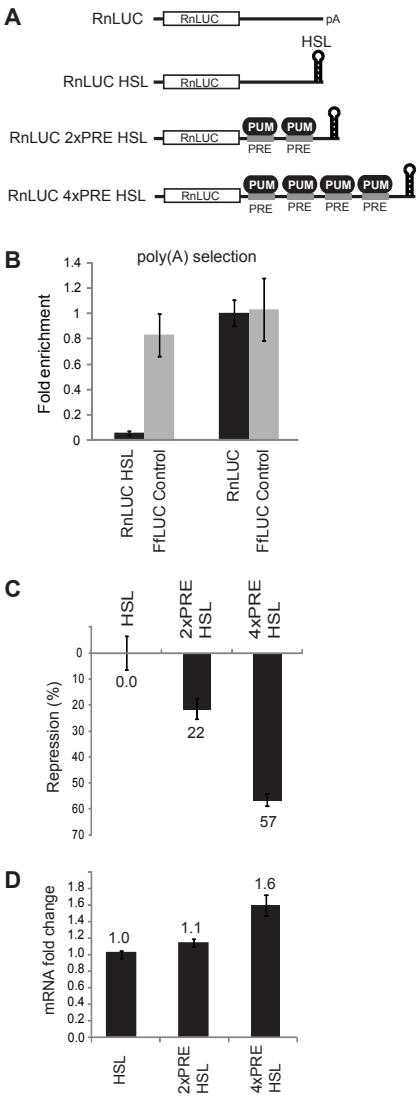


Figure 3.7. Poly(A) independent repression by PUMs.

A. *Renilla* luciferase reporters (RnLUC) that lack a 3' poly(A) tail were created by replacing the cleavage/poly-adenylation sites with a histone stem loop (HSL) processing signal. Two or four PREs were inserted into the 3'UTR to created RnLUC 2xPRE HSL and RnLUC 4xPRE HSL, respectively. *B.* Graph of fold enrichment of RnLUC HSL, RnLUC, and FfLUC internal control mRNAs in poly(A) selected fraction isolated using oligo-dT affinity purification. Fold enrichment was measured by qRT-PCR analysis of poly(A) selected mRNA, normalized to total, and calculated relative to polyadenylated RnLUC. *C*. Graph of percent repression relative to RnLUC HSL for the indicated reporters showing that endogenous PUMs repress 2x and 4xPRE HSL reporters. *D*. Graph of fold change in reporter mRNA levels measured by multiplexed qRT-PCR and calculated relative to RnLUC HSL control.

CHAPTER 4

Identification of mRNAs regulated by human Pumilio proteins

This work was carried out in collaboration with Trista Schagat, Rich McEachin, and Ashwini Bhasi and is currently in preparation for publication. Schagat carried out PUM knockdown in HEK293 cells, RNA isolation, quality control, and library preparation. McEachin and Bhasi carried out analysis of RNA Seq raw data. Van Etten validated RNA Seq targets by qRT-PCR, and performed additional bioinformatics analyses including GO analysis.

Introduction

Regulation of gene expression is a complex and highly regulated process: failure to properly control translation or decay of messenger RNAs results in aberrant gene expression and disease. PUF (Pumilio and FBF) proteins are a family of conserved eukaryotic RNA binding repressors that bind to specific sequences in the 3' UTR of messenger RNAs (1). All PUFs are characterized by the presence of an RNA binding domain (RBD) that contains 8 α-helical modular repeats, each of which binds a single

ribonucleotide (2-7). The binding interaction between each repeat and RNA nucleotide is mediated by three amino acids that recognize and confer binding specificity (7).

PUFs control many biological processes in model organisms including: stem cell proliferation, fertility, development, and learning and memory, (3, 4, 8-25). Mammalian PUFs are thought to play a role in cell proliferation, neural function, and fertility (24-30).

PUFs are generally known to behave as repressors in model organisms(1, 31). It is largely understood that PUFs in model organisms promote mRNA decay: yeast PUFs promote decay through deadenylation, enzymatic removal of the poly (A) tail, and decapping, enzymatic removal of the 5' cap (32-34). *Drosophila* and *C. elegans* PUF repression correlates with deadenylation (8, 22, 35-37). PUFs are thought to act in concert with a number of co-repressor proteins in order to control the stability and translation of target mRNAs (4, 22, 30, 32-36, 38-42). In yeast, PUFs enhance deadenylation by binding to the Pop2p deadenylase subunit to promote deadenylation by the deadenylase Ccr4p (32, 33). The interaction between PUFs and POP2 orthologs is conserved from yeast to human(33, 41). Since mRNA decay and translational control are intimately linked, downstream effects of PUF mediated mRNA decay include translational repression, though at least one study found that yeast and C. elegans PUF cause direct inhibition of translation (39).

Humans have two canonical PUF proteins, called PUM1 and PUM2, which are highly similar and expressed widely among tissues. Among their conserved RNA binding domains they share 91% identity and outside that region they share 75% identity (43, 44). PUM1 and PUM2 bind with high affinity to the sequence UGUANAUA: we will refer to this sequence as the PUM response element (PRE) (45, 46). Human PUM1 and PUM2

cause repression of reporter mRNAs by deadenylation-dependent and independent mechanisms (41). Overexpression of PUM2 with NANOS3 inhibits PRE-containing E2F3 reporters in TCCSUP bladder carcinoma cells and the authors argue that microRNA mediated repression in this case is facilitated by PUMs (29). PUM1 was reported to repress the p27 tumor suppressor gene (25). Recent studies indicate that they bind to deadenylases and promote decay of target mRNAs (41, 46). Furthermore, overexpression of PUM2 represses ERK2 and p38α reporter genes in human embryonic stem cells (47). Together with work done in model organisms, these studies suggest that that human PUMs enact evolutionarily conserved mechanisms to promote decay of mRNA targets.

While it is known that PUM1 and PUM2 cause mRNA decay and protein repression, the identities of their targets are less clear (41). RNP immunoprecipitation-microarray (RIP-Chip), studies indicate that both human PUM1 and PUM2 bind hundreds of mRNAs, yet this binding data remains largely unexplored with functional assays (28, 45, 46). In one instance, depletion of PUM1 by RNAi modestly stabilized five mRNAs (46). PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation) and RIP-Chip (RNP immunoprecipitation –microarray) studies identified RNAs bound to PUM1 and PUM2. RIP-Chip and PAR-CLIP techniques rely on high quality antibodies to coimmunoprecipitate bound RNAs and provide detailed information about the identities of the RNAs bound. RIP-Chip employs microarray analysis to identify bound RNAs to RBPs: hundreds of mRNAs bound to human PUM1 and PUM2 were identified using this technique (45, 46). PAR-CLIP relies on photactivation of incorporated 4-thiouridine (4SU) into RNA transcripts in cells (48).

Associated RBPs are then crosslinked to 4SU-labeled transcripts and immunoprecipitated from cultured cells, and subjected to deep sequencing (48). This technique enables high throughput identification of RNAs bound to specific proteins. Studies employing this technique identified 2847 bound RNA targets of PUM2-FLAG fusion proteins in HEK293 cells (48).

We previously reported robust depletion of both PUM1 and PUM2. We also demonstrated that knockdown significantly de-repressed PUM target mRNAs. It is important to note that, since both PUM1 and PUM2 are potent repressors of reporter mRNAs containing a PRE, their functions may overlap in specific instances (41). Therefore, studies measuring regulation by either PUM1 or PUM2 alone may miss important regulatory information.

Here we seek to identify target mRNAs repressed by PUM1 and PUM2 using RNAi of *both* PUM1 and PUM2 in human cells coupled with comparative RNA Sequencing (RNA-Seq) analysis between control and knockdown samples. RNA-Seq, or whole transcriptome shotgun sequencing, is a methodology that uses next generation sequencing to identify RNAs present across a broad dynamic range in a sequence library. This technique can therefore be used to detect changes in expression level among thousands of RNAs in response to depletion of PUM1 and PUM2. The advantages of RNA-Seq analysis are numerous: most importantly, its dynamic range is broad and quantitative (49). Furthermore, it is possible to carry out multiple replicates in one experiment. By knocking down PUM1 and PUM2 in HEK293 cells, we circumvent possible problems with the proteins' potential overlapping functionality.

Few mRNAs that are directly regulated by PUMs have been reported; therefore, it is essential in order to understand the biological functions of PUMs. We sought to identify and validate mRNAs that are significantly regulated upon knockdown of PUMs in HEK293 cells. We took a multidimensional approach to identifying PUM regulated target mRNAs using the following assumptions: 1) Abundance of PUM target mRNAs should reproducibly change whenever the PUM is depleted. 2) PUM target mRNAs should contain a PRE motif. The binding affinity and specificity of PUM1 and PUM2 has been extensively analyzed and the PRE consensus corresponds to UGUANAUA(45). We previously showed that the presence of the PRE in an mRNA can elicit up to a 4-fold reduction in target mRNA abundance (41). To identify high confidence target mRNAs, we then compared those that are differentially regulated by PUMs to those that contain PREs and/or were experimentally shown to bind to PUMs in RIP-CHIP and PAR-CLIP studies (45, 46, 48). We identified over 1000 RNAs that were differentially expressed upon PUM knockdown and found that 486 of those RNAs contain predicted PREs. We found 287 that met the following criteria: 1) The RNA was differentially expressed in our dataset, 2) The RNA contains at least one PRE, and 3) The RNA was identified in previous RIP-Chip or PAR-CLIP analyses. Furthermore, we went on to demonstrate validation of multiple PUM targets by qRT-PCR analysis. Together, our results provide insight into the potential regulatory roles of PUMs in biological processes. Using highly sensitive RNA-Seq analysis, we identify a set of high confidence direct PUM RNA targets and, importantly, take into account the presence of both PUM1 and PUM2 in human cells.

Materials and methods

Cell culture and siRNA transfections

Human HEK293 cells were cultured at 37°C under 5% CO_2 in DMEM with glucose and 1 x Penicillin/Streptomycin/Glutamine and 10% FBS (Gibco) as described previously (41). PUMs were knocked down in HEK293 cells using On-target Plus Smartpool siRNAs for PUM1 (L-014179-00), PUM2 (L-014031-02), or non-targeting control siRNAs (Dharmacon). For RNA Seq experiments, HEK293 cells (2 x 10^5 cells per well) were plated into a 12-well plate. After 24 hours, 500 µL culture medium was removed and cells were transfected with 100 fmoles of siRNAs using Dharmafect-I (Dharmacon). After 4-8 hours of siRNA treatment, 500 µL culture medium was refreshed. After 24 hours, 500 µL culture medium was removed and cells were transfected again with 100 fmoles of siRNAs using Dharmafect-I. After 4-8 hours of siRNA treatment, 650 µL culture medium was refreshed. Cells were harvested 48 hours after last transfection for RNA purification and RNA Seq experiments.

Western blotting

Protein lysates were created by lysing cells on ice in TNEMN-150 plus 1X protease inhibitor cocktail as described in (41). SDS PAGE loading buffer was added to samples to a final concentration of 1X and the samples were boiled for 10 minutes. Lysates were run on a 12% SDSPAGE gel for 1.5-2 hours at 100 volts and transferred at 80V for 2 hours at 4°C. Blots were blocked in blotto (5%: 5g dried milk in 100 mL 1X PBS and 0.1% Tween-20) overnight at 4C. Blots were incubated in primary antibody for 1 hour at room temperature with gentle rocking, washed 3 times for 10 minutes with blotto, and incubated in secondary antibody for 1 hour at room temperature with gentle

rocking. Blots were washed 3 times for 10 minutes with gentle rocking, developed with Millipore Immobilon reagent, and visualized with X-ray film. Antibodies used in these experiments were described in (41).

RNA Seq

Control samples were transfected in triplicate with nontargeting siRNA pools as described above. PUM knockdown samples were transfected in triplicate with PUM1 and PUM2 siRNA smartpools as described above. RNA-Seq analysis was performed on poly (A) enriched RNA isolated from HEK293 cells. Total RNA was purified using the Maxwell simplyRNA cells kit (Promega). Polyadenylated RNAs were isolated using oligo dT purification and prepared for sequencing according to manufacturer instructions for Illumina sequencing. Bar coded libraries for each sample were generated according to the Illumina TruSeq kit instructions. Purity and integrity of total RNA was using a Bioanalyzer before deep sequencing. All total RNA samples had RIN values of 10, which is the highest possible value, and indicates that they are high quality nucleic acid samples.

Samples were submitted to the University of Michigan Sequencing Core for Illumina sequencing, paired end, 100 base sequencing on an Illumina HiSeq instrument. Libraries were multiplexed in a single "lane". More than 26 million sequence reads per sample were obtained for each of six samples: 3 control NTC siRNA and 3 PUM1/2 knockdown. Low quality sequence reads were filtered by Illumina's CASAVA software. The resulting dataset was subjected to FASTQC analysis and trimmed to give 86 good quality base reads.

The data analysis pipeline included analysis with the Tuxedo suite: 1) TopHat and Bowtie were used to map and align reads, Cufflinks was then used to quantify abundance of each transcript, and Cuffmerge and Cuffdiff (2.1) using a false discovery rate of ≤ 0.05 were used to assess differential abundance as described in (49). Relative abundance was reported in reads per kilobase of exon per million fragments (RPKM values).

RNA purification and cDNA amplification with reverse transcriptase

Three replicates of cell populations transfected with either nontargeting control siRNA or PUM1/2 siRNA smartpools were harvested for RNA purification and qRT-PCR analysis. RNA was purified from HEK293 cells in 6 well dishes after siRNA treatment using the Maxwell RNA simplyRNA cells kit as described previously, including the on bead DNase treatment (41). cDNA was synthesized using GoScript Reverse Transcriptase as indicated in manufacturer instructions and as described in (41). After cDNA synthesis, samples were diluted appropriately in nuclease free water and used for qPCR validation.

Quantitative PCR

qPCR was carried out using GoTaq qPCR (Promega). Cycling conditions were as follows: (i) 95°C for 3 minutes, (ii) 95°C for 10 sec, (iii) 65°C for 30 sec, and (iv) 72°C for 40 sec. Steps (ii) through (iv) were repeated for a total of 40 cycles. Negative control reactions were performed in the absence of template or reverse transcriptase. Cycle thresholds (Ct) were measured using the CFX Manager software and analyzed using the ΔΔCt method (50, 51). ΔCt was calculated by normalizing to the 18S gene Ct values. We then calculated ΔΔCt as follows: ΔΔCt= ΔCt(target RNAi)- ΔCt (control RNAi). All

primer set amplification efficiencies were optimized to 90-110% at 200 nM final concentration. qPCR primer sequences are as follows:

ANO4 FWD: GAAGTACTTGATTCAGGAGATGATGTATG
ANO4 REV: GAAGAAGTAAACAGAATTCAGGTCCTTGG
NOVA2 FWD: CTCATCAGTCAGCGGGTCACCTAC
NOVA2 REV: GAGGAGCAGGACTACACCAAGCTG
DEK FWD: GCAAAAAGGTCTATGAAAATTATCCTACTTATG
DEK REV: GAAAAGGAAATACATTCTCTTTGCTGG
RET FWD: CTCGAGCCCTCCCTTCCACAT
RET REV: AAGCATCCAGTTAGCATATACACTATCATTTG
DUSP6 FWD: ACAACAGGGTTCCAGCACAG
DUSP6 REV: AAACTGCTGAAGGGCCAGAC
ETV4 FWD: CTGCGACCATTCCCAGATGATGTC
ETV4 REV: CAAGGCCACCAGAAATTGCCAC
L1CAM FWD: GGCGGCAAATACTCAGTGAAGG
L1CAM REV: TACTCGCCGAAGGTCTCATC
SCUBE1 FWD: GATCCGAGATGCCAAGTGCCA
SCUBE1 REV: CGATCTCAAACTCTGCTGTG
SMPDL3A FWD: CAAATCTCCAGGTTTTCCCTGCGC
SMPDL3A REV: GAATAAAAACCACCTTTCCTTAAAGTACTAATAGC
FZD8 FWD: GGCGAGCTCCGTGTCTTATCC
FZD8 REV: CTTCGCTGCACTTGGCTCTCCT
FMR1 FWD: CCGAACAGATAATCGTCCACGTAA
FMR1 REV: GTGCGCAGCCGACTACCTTCA
PUM1 FWD: CATGAACGACGGTCCCCACAG
PUM1 REV: GCTTGGCCAGAATGTGCTTGC
PUM2 FWD: ATGCAATACTGTCTCCGCGATCAG
PUM2 REV: CTTTCTCAGGTCCATCTGTTTCAGC
18S FWD: CAGCCACCCGAGATTGAGCA
18S REV: TAGTAGCGACGGGCGGTGTG

Results

PUM1 and PUM2 regulate the abundance of more than 1000 mRNAs

PUM1 and PUM2 were depleted from HEK293 cells by RNA interference mediated knockdown with specific siRNAs. Triplicate cell populations were treated with PUM1 and PUM2 siRNAs (PUM1/2 kd). Triplicate cell populations were treated with

non-targeting control siRNAs (NTC) as a negative control. Depletion of PUM1 and PUM2 in the PUM1/2 siRNA treated cells was confirmed by Western blot. Neither PUM1 nor PUM2 protein was detected in extract from cells treated with the PUM1 and PUM2 siRNAs, whereas both were readily detected in NTC extracts (Fig. 4.1A). Additionally, we previously demonstrated that RNAi depletion of PUM1 and PUM2 using this method specifically alleviated PUM repression of a luciferase reporter containing PREs in its 3' UTR (Fig. 3.1) (41).

Total RNA was purified from each cell population, and the integrity of each RNA sample was verified by Bioanalyzer analysis. Next, poly-adenylated RNAs were purified from each sample using oligo-dT affinity purification. Barcoded cDNA libraries were generated for each sample according to Illumina Tru-Seq library generation method. The workflow for sample preparation and library generation is outlined in Figure 4.1B. Sequencing was performed with a HiSeq system (Illumina). More than 26 million reads per sample were obtained for each sample and corresponding to more than 18500 genes. Using this data, normalized expression values were determined for each gene in units of RPKM (reads per kilobase per million). The RPKM values spanned a range of 7 orders of magnitude (Range of 0.0009-1788) (Table 4.1, Supplemental table S1).

The resulting RNA abundance data was analyzed to detect statistically significant changes in abundance between NTC and PUM1+2 kd conditions. 1035 genes were reproducibly, differentially expressed in the PUM1+2 knockdown sample (Table 4.1). The complete list of differentially expressed targets is available in the supplemental information (Table S1). Fold changes of these mRNAs ranged from 3.81 to 0.287, relative to the NTC control. We note that fold changes with larger magnitudes were

observed but these did not meet criteria for statistical significance. As expected, PUM1 and PUM2 mRNAs were present in the differentially regulated dataset and confirmed that they were depleted by corresponding siRNAs by 64% and 73%, respectively.

Of the 1035 differentially regulated genes, 695 RNAs increased in abundance in the absence of PUMs, indicating that they are negatively regulated by PUM1 and PUM2 (Fig. 4.2A, Table S1 in red). Conversely, 340 RNAs were reduced in abundance upon PUM knockdown, implicating PUMs in positive regulation those genes (Figure 4.2B, Table S1, highlighted in green). One basic question that remained unanswered is: do PUM target RNAs fall within a specific expression range? Are PUM targets abundant or rare? The results shown in Fig. 4.2 demonstrate that PUM regulated RNAs span all expression ranges detected: therefore, PUMs do not appear to preferentially regulate genes in a manner that depends on their expression level (Fig .4.2A, B).

By imposing an arbitrary cutoff of ≤1.5 –fold or ≥1.5-fold for differential expression: we find that 387 genes are affected by knockdown of PUM1 and PUM2. These genes included 271 that increased in abundance whereas 116 of the RNAs decreased in abundance. Two mRNAs meeting this cutoff were previously identified as putative PUM1 targets in RIP-Chip studies, and were validated by qRT-PCR (45, 46). We next asked if PUM mRNA targets previously reported in the literature were differentially regulated in our dataset. The SLBP mRNA, which contains 1 3' UTR PRE, was reported to be stabilized upon PUM1 knockdown (46). We corroborated its regulation in our dataset, and found that it is upregulated 1.55-fold upon PUM1/2 knockdown. We also discovered that PCNA mRNA, which contains 2 PREs in its 3' UTR, was upregulated 1.56-fold upon PUM knockdown. PCNA mRNA encodes a

nuclear protein involved in replication and was reported to be modestly stabilized by PUM1 knockdown in a previous study (46). Two other PRE containing mRNAs, CKS2 and UBA2, were upregulated 1.48 and 1.41-fold, respectively, upon PUM1/2 knockdown. Both mRNAs contain PREs in their 3' UTRs. UBA2 mRNA encodes sumoylation ligase that is involved in p53 induced apoptotic activity and was identified in a RIP-Chip study examining PUF regulated mRNAs in PUM1 deficient mice: UBA2 mRNA isolated from PUM1 deficient mice was upregulated in qRT-PCR studies (28). The CKS2 mRNA encodes a protein that binds the catalytic subunit of cyclin dependent kinases and is cell cycle regulated (46). Several mRNAs that were previously verified in RIP ChIP studies including p27, ERK-2, p38 α , and E2F3, did not meet statistical tests in our study (25, 29, 47).

We next sought to gain insight into the physiological functions of genes encoded by mRNAs differentially regulated in our dataset. Gene ontology analysis was carried out using PANTHER analysis for each of three GO terms, biological process, molecular function, and cellular component, on our dataset of 1033 differentially regulated PUM target mRNAs (52, 53). Results from this study are shown in Table 4.2. Note that many genes fell into the "unclassified" GO term, and were thusly removed from the table. Several GO terms from three categories, Biological Process, Cellular Component, and Molecular Function, were enriched in our dataset (Table 4.2). Many genes in the dataset are involved in developmental processes, which is expected given what is known about PUF function in model organisms; here, we see that developmental process, system development, nervous system development, ectoderm development, and mesoderm development are statistically overrepresented in our dataset (Table 4.2). Furthermore, cell

communication, cell surface receptor linked signal transduction, and cell adhesion are also overrepresented in the dataset (Table 4.2).

KEGG pathway analysis was carried out on the same list of mRNA targets using the DAVID functional annotation tool (Table 4.3) (54-56). Several pathways are enriched in our dataset including several cardiomyopathies, MAPK, cytokine, TGF-beta, and p53 signaling pathways, as well as regulation of the actin cytoskeleton and melanogenesis (Table 4.3). It is important to note that axon guidance is enriched 2.59-fold (p-value $1.29x10^{-4}$) in our dataset (Table 4.3). There is a great deal of work implicating *Drosophila* PUM in neuronal function and axon guidance, as well as in mammalian neurons; therefore, human PUMs may serve similar regulatory roles in the nervous system (16, 17, 21, 57-59). We also discovered that p53 signaling pathway is overrepresented in our dataset 2.34-fold (p-value $2.5x10^{-2}$): here, we show that PUMs affect expression of downstream targets of p53 signaling including CDKN1A (p21), SFN, Lrdd, BBC3, Cyclin D, BAI-1, RRM2B, Cyclin G, Cdc2, and CASP3 (Table 4.3). Further, it was shown that mammalian PUM1 represses regulators of the p53 pathway in mice (28). These analyses indicate that PUMs regulate multiple components of several pathways including cell proliferation pathways, tumor suppressor signaling, and neural function.

Transcriptome-wide prediction of human PUM response elements

We next sought to determine the number of RNAs that contain predicted PUM response elements (PREs). We did so in order to extract the highest probability targets from our differentially regulated dataset. Direct targets of PUM proteins should contain at least one PRE; therefore, we selected only those RNAs for further analysis.

We limited our search to the well-characterized, high affinity PRE consensus site UGUANAUA, which has been shown to be necessary and sufficient to confer PUM repression of a reporter mRNA (41). We compared differentially regulated genes in our dataset against the human transcriptome (UCSC genome version Hg19), which contains annotations for 40,901 RNAs, including 34,375 mRNAs and 6526 ncRNAs. Of those annotated RNAs, 9304 mRNAs had UGUANAUA PRE sites. This set of RNAs included isoforms produced by the same gene. When condensed, 5235 unique genes contained one or more predicted PRE motifs. 80.2% of predicted PREs are located in the 3'UTR. It is noteworthy that nearly all known PUF regulated mRNAs have PUF binding sites present in the 3'UTR. 16.4% of predicted PREs reside in coding sequences and only 2.0% of PREs of predicted PREs are located in the 5' UTR of transcripts (Fig. 4.3A). Several genes contain PREs in multiple locations along the transcript: 67 genes contain at least one site in each untranslated region, and 7 genes contain at least one in the 5' UTR, coding region, and 3' UTR (Fig. 4.3A). Noncoding RNAs were discovered to contain PREs, sometimes in large numbers depending on the transcript: 764 of the total number of PREs identified (Fig. 4.3A). The discovery of polyadenylated noncoding in the differentially regulated dataset is surprising: there are no documented cases of PUF-mediated regulation of a noncoding RNA to date.

Known PUF targets contain PREs in their 3' UTRs, however, it has yet to be disproven that PREs are located elsewhere in the mRNA. Intriguingly, many mRNAs or ncRNAs contain more than one PRE. The average number of predicted PREs per transcript is ~ 1.5, however, the range is between 1 and 16 PREs in mRNAs and ncRNAs. RNAs are predicted to contain between 1 and 13 PREs in the 3' UTR, 1-2 in the 5' UTR,

and between 1-10 in the coding region (Figure 4.3B). Several ncRNAs identified in our dataset contain multiple predicted PUM sites, ranging between 1 per transcript and 16 in some RNAs: specifically, the ncRNA HCG11 contains 16 predicted PREs, and LOC647979 contains 15 predicted PREs.

PREs are enriched in the 3' UTR of target mRNAs

To further narrow our list of strong direct PUM targets, we determined which RNAs contained PREs using our prediction analysis. We found that 486 PUM target RNAs, including ncRNAs and mRNAs, contain at least one PRE. 406 PRE-containing RNAs were upregulated in response to PUM knockdown, consistent with these being repressed by PUMs. The remaining 80 of PRE-containing RNAs, were downregulated upon PUM knockdown, which suggests that they are activated directly or indirectly in response to PUM activity.

Next we sought to determine the locations of PREs in differentially regulated PUM target genes. Of those predicted sites found in upregulated mRNAs, 1199 PREs were found in the 3'UTRs of the 406 strong direct PUM targets we identified. Furthermore, 108 PREs are located in the ORF, and 6 PREs are in the 5'UTR. 214 predicted PREs were found in the 80 genes potentially activated by PUMs: 189 sites are located in the 3'UTR, 17 are in the ORF, and 8 are in the 5'UTR. From these data, we conclude that mRNAs responsive to PUM regulation have PREs primarily in their 3' UTRs. The RAD51AP1 mRNA contains one PRE in its 5' UTRs and is upregulated in our differentially expressed dataset 1.3-fold. Furthermore, 12 upregulated mRNAs contain one PRE in their ORF but not in either UTR, including: BOD1L, COL21A1, DOCK9, DSEL, FBN2, GMCL1, KDR, NBEAL1, NCKAP5, NTS, SLC16A9, and

ZNF711. Four downregulated mRNAs including: FHL1, IFIT2, STAT1, and STOX2 have a PRE only in their 5' UTR. Ten mRNAs including: APPBP2, ATL3, CBWD1, CTDNEP1, IFIT1, IFIT3, NOL11, PKDREJ, PPP2R2A, and PPP2R2D are downregulated in our differentially expressed dataset and contain PREs in their ORFs only. Since a small percentage of PREs are located in the ORF and 5'UTR of target RNAs, the possibility that regulation occurs through sites located elsewhere in the mRNA cannot be ruled out.

Identification of high confidence PUM RNA targets

Our next goal was to compare differentially expressed RNA targets that contain PREs with PUM bound (or associated) RNAs identified in RIP ChIP and PAR-CLIP studies (45, 46, 48). Morris et al. used RIP-Chip to identify PUM1 associated RNAs in HeLa S3 cells: 726 associated RNAs were identified in this study and several genes were validated by qRT-PCR (46). We know from previous work that PUMs have overlapping functions in human cells, and therefore, it is likely that many regulated PUM targets were omitted from the Morris study, as PUM2 was not considered for analysis (41). Galgano et al. performed RIP-Chip with antibodies to both human PUM1 and PUM2 in HeLa S3 cells and thusly identified 1040 PUM1 and 435 PUM2 associated RNAs, about half of which contain PREs (45). Hafner et al. performed PAR-CLIP to identify PUM2 associated RNAs and identified ~3000 transcripts associated with PUM2 isolated from HEK293 cells (48). Together, these analyses provide a list of PUM1 and PUM2 bound mRNAs. Experimental limitations imposed by PAR-CLIP and RIP-Chip, such as expression and detection of mRNAs, quality of antibodies used for immunopurification, and crosslinking efficiency, likely leave many associated RNAs from our purview.

The purpose of our analysis is to test whether or not previously established evidence of widespread PUM regulation and their targets agree with our own predictions and findings in order to generate a set of high confidence PUM target genes. First we identified those genes that met criteria for differential expression: this is called the "Response" dataset, and consists of 1033 genes either up- or down-regulated in response to PUM knockdown. We compared those to 5235 genes predicted to contain at least one PRE, called the "Predicted PRE" dataset, and found that 486 genes meet both sets of criteria (Fig. 4.4A). Together, RIP ChIP and PAR-CLIP studied yielded a set of 4071 unique genes that bind either human PUM1, PUM2, or both: this will be referred to as the "Bound" dataset. We next compared the genes found in the Bound dataset to the 1033 genes in the Response dataset. Comparison between Bound and Response yielded 389 overlapping genes between our RNA Seq analysis and RIP-Chip and PAR-CLIP studies (Fig. 4.4B). We also compared the genes found in the Bound data with the Predicted PRE dataset and identified 2007 genes of 4071 Bound RNAs that met both criteria (Fig. 4.6C). Lastly, we compared all three studies to identify the highest confidence PUM targets: 287 genes contain at least one predicted PRE, were discovered in RIP-Chip and PAR-CLIP assays to bind either PUM1 or PUM2, and met criteria for differential regulation in our RNA Seq studies (Fig. 4.4C).

Biological significance of newly identified PUM target RNAs

We next wanted to determine if PUMs control specific pathways. To do so, we performed gene ontology analyses using DAVID and PANTHER on high quality PUM target RNAs. We honed in on those target RNAs that are differentially regulated in our dataset. Our analyses in the following sections will focus on key features and known

roles of RNAs in the Response dataset that contain predicted PUM sites in order to gain a better understanding of their roles in biology and how PUM regulation might affect physiological processes.

After identifying many hundreds of activated and repressed PUM targets, we sought to better understand their functional roles in human cells. We performed gene ontology analyses on each upregulated and downregulated dataset independently with the intention of finding specific pathways that are regulated by PUMs. Through these analyses we discovered that PUMs regulate genes encoding components of multiple biological processes including cell communication, nervous system, mesoderm, heart, and ectoderm development, as well as cell adhesion were statistically enriched with p-values ≤ 0.05 (Table. 4.4). Interestingly, repressed PUM targets are statistically enriched in extracellular matrix categories, and appear to regulate genes involved in receptor activity and binding, and enzyme regulation (Table 4.4). KEGG pathway analysis indicated PUM regulated RNAs are involved in axon guidance, p53 signaling, and several other pathways (Table 4.3). We found that, of those RNAs identified in GO KEGG pathway analysis for the p53 signaling pathway, six contain predicted PREs including p21 (CDKN1A), p53R2 (RRM2B), PUMA (BBC3), Cyclin G (CCNG2), CDC2 (CDK1), and CASP3. This data supports that PUMs repress multiple components of biological pathways.

We analyzed PUM roles in disease pathways using DAVID and discovered that repressed PUM targets are statistically enriched in disease pathways, which include disorders of lipid metabolism, as well as psychiatric conditions and developmental disorders (Figure S2). PANTHER analysis yielded no results with p-values ≤ 0.05 in

terms of statistical enrichment of activated PUM targets; intriguingly, however, activated targets were enriched in intracellular and cytoskeleton cellular components (Supplemental Fig. S2).

Secondary validation of human PUM regulated mRNAs

We next sought to validate regulation of selected high confidence PUM mRNA targets. mRNAs were chosen based on their relevance to human disease, as well as their potential for PUM regulation (ie: they exhibited high fold change among differentially regulated genes). Expression of the following repressed PUM target mRNAs: SCUBE1, SMPLD3A, DEK, ANO4, FMR1, FZD8, L1CAM, NOVA2, and RET. Furthermore, the following activated PUM targets were analyzed: DUSP6, and ETV4. Knockdown of PUMs was performed in HEK293 cells in triplicate following the same protocol that was used to prepare RNA-Seq cDNA libraries: cells were then harvested for RNA purification and cDNA synthesis with reverse transcriptase.

Next, we quantified changes between control and PUM k/d samples using quantitative PCR analysis: duplicate qPCR assays were carried out for each of three RNA samples for either PUM kd or NTC. We anticipated that relative changes in RNA level in qPCR assays would mirror changes seen in RNA-Seq analysis. Relative fold changes were calculated using the comparative Ct method by which we normalized each sample to the internal control 18S ribosomal RNA. We discovered that each selected gene was either increased or decreased upon knockdown of PUM1 and PUM2 to nearly the same degree as in RNA Seq experiments (Table 4.5). We wished to identify targets involved in key regulatory pathways and in human disease, and therefore chose a number of oncogenes and neural regulators to validate by qRT-PCR (Table 4.6). Additionally, all

validated targets contain at least one PRE, but several have multiple PREs (Table 4.7). Most sites lie in the 3'UTR of target mRNAs, however, the DEK oncogene contains three PREs, two of which are located in the 3' UTR, but one is located in the ORF. It is unknown if the PRE located in the ORF confers regulation, as no PUM targets confirmed to date contain ORF or 5' UTR PREs.

DEK is an oncogene whose upregulation is associated with many cancers. Upon depletion of PUM1 and PUM2, DEK mRNA increased in abundance by 1.6 fold in theRNA Seq assay and 1.2-fold in the qRT-PCR assay (Table 4.5). Another PUM target, the mRNA encoding FMR1, deficiency of FMR1 causes Fragile-X mental retardation and premature ovarian failure in female patients is upregulated 1.3-fold in RNA Seq experiments and 1.4-fold in qPCR studies (Table 4.5). Two highly regulated mRNAs encode SCUBE1 and SMPDL3A, which are two proteins thought to be involved in vascular biology and bladder cancer, respectively. SCUBE1 is a marker of platelet activation and is thought to be upregulated in hypertension(60). SCUBE1 mRNA increases by 2.4-fold (RNA-Seq) and 2.2-fold (qPCR) upon PUM knockdown (Table 4.5). SMPDL3A is characterized as a binding partner of DBCCR1 (deleted in bladder cancer 1), which when overexpressed, leads to upregulation of SMPDL3A (61). SMPDL3A mRNA increased 2.4-fold (RNA Seq) and 2.3-fold (qPCR) upon PUM knockdown from HEK293 cells (Table 4.5).

Intriguingly, several high confidence PUM target mRNAs that we identified decreased in abundance upon PUM knockdown, which indicates that they may be directly activated by PUMs (Table 4.5). DUSP6 is a cytoplasmic dual-specificity phosphatase that is thought to target two MAP kinases, ERK1 and ERK2 (62). Our

analysis demonstrates that this mRNA decreased 0.5-fold in RNA Seq and 0.7-fold in qPCR analyses upon PUM knockdown (Table 4.5). ETV4, a transcription factor, decreased markedly in our dataset by 0.3-fold (RNA Seq), and 0.4-fold in qPCR assays (Table 4.5). Further study will be necessary to dissect the regulatory role human PUMs play in the control of these biological target mRNAs, and our results demonstrate their status as PUM targets and provide researchers a starting point for future mechanistic study.

Discussion

We identified 1033 RNAs that were differentially regulated when PUMs were depleted with siRNAs in HEK293 cells. 486 of those differentially regulated RNAs contain at least one PRE and were considered high quality PUM targets. That is not to say that targets that do not contain a PRE are not subjected to PUM-mediated regulation, however regulation of targets lacking a PRE are likely indirect targets, and are controlled by an intermediary factor affected directly by PUMs. Furthermore, we recognize the potential for false negatives in our experimental design: the cell population and subsequently purified RNA population was not a null PUM background, samples with too much variability between replicates were ignored as they did not pass statistical cutoffs, and many genes were not detected or were too low in this cell population. We detected four RNAs in our dataset that were previously identified as PUM targets, including SLBP, PCNA, CKS2, and UBA2. It is likely that we did not identify other targets, including many p53 regulators, MAP kinases, and p27, in our dataset because of limitations of our RNA-Seq experiment. RNA targets may have been missed because, under specific conditions, they have shortened 3' UTRs lacking PRE elements necessary

for PUM repression. We are confident that those RNAs we identified that contain at least one PRE are high confidence PUM targets, but recognize that there may be many more RNA targets of PUMs that contain weaker PREs that could be responsive to regulation.

Among those targets containing at least one PRE, 287 RNAs were associated with PUM1 or PUM2 in previous analyses (45, 46, 48). Gene ontological analyses indicate that several biological pathways are represented among our dataset of PUM targets that contain PREs, including developmental processes, tumor suppressor signaling pathways, and neural pathways. From these data, we conclude that both PUM1 and PUM2 are regulators of many RNAs involved in multiple pathways.

486 RNA targets that contain at least 1 PRE are differentially expressed in HEK293 cells depleted of PUM1 and PUM2. Importantly, this is the first study that looks directly at levels of RNA in response to changes in endogenous PUM1 and PUM2 protein levels in human cells. PREs are enriched in 3' UTRs of high confidence targets we identified: it will be interesting in the future to test genes that contain PREs in several sites: this could provide insight into a new mode of PUF regulation, or possibly a mechanism by which PUF repression is strengthened in a combinatorial fashion.

We compared RNA Seq, PRE prediction, gene ontology analyses, and comparison with other datasets to provide a high quality set of RNA targets of PUM regulation and demonstrate the tremendous potential for PUM control of many aspects of normal and abnormal physiological processes in humans. 406 targets in our dataset are strong candidates for PUM mediated repression, as they increase in abundance upon PUM knockdown and contain at least one PRE. We know that both PUMs are potent repressors of a reporter RNA containing 3 PREs in its 3' UTR via deadenylation-

dependent and independent means; therefore, our data suggest that these 406 genes are direct targets of PUMs but functional validation and PUMs' effects on their decay remains to be shown (41). Regulation of these 406 RNAs likely impacts a breadth of biological processes and future work will focus on validating PUM regulation of RNAs involved specific pathways. Gene ontology analysis and previous studies suggest involvement of mammalian PUMs in nervous system function, development, and cell proliferation; therefore, it will be of future interest to focus on PUM regulation of these pathways.

We were surprised to find that 80 PRE-containing RNAs were decreased in abundance upon PUM knockdown and are therefore candidates for activation by PUMs. PUM target activation has been demonstrated in few studies. This could occur by a number of mechanisms: PUMs may repress a repressor, which would in turn, result in upregulation of an indirect PUM target RNA. Since we find that 80 activated PUM target RNAs contain PRE, however, it is more likely that a direct activation mechanism exists to control the abundance of these RNAs. PUMs may carry out activation through a mechanism that is yet to be discovered: similar to their repressive mechanisms, PUMs may bind to the RNA through a specific interaction with the PRE, but instead recruit RBPs or effectors that activate their targets. Alternatively, other *cis* sequence elements may contribute to combinatorial control by several different RBPs including PUMs to activate RNA targets. Insertion of a PRE near a CPE (cytoplasmic polyadenylation element) results in a 2-fold enhancement of translation due to stabilization of cytoplasmic polyadenylation element binding protein (CPEB) on the mRNA (63). Furthermore, upon progesterone treatment in oocytes, PUM2 dissociation from the RINGO/Spy mRNA

results in translation of RINGO/Spy mRNA, which is required for CPEB activation and cytoplasmic polyadenylation of downstream targets (40). Therefore, combinatorial control by a number of RBPs to activate PUM target RNAs is a possible mechanism for activation.

We validated eleven mRNA targets by qRT-PCR. We chose genes that represent a broad range of RNAs encoding proteins involved in numerous physiological functions. The potential for PUM regulation in various physiological processes and disease pathways is broad: for example, if PUMs are hyperactive or underactive in specific tissues, disease may ensue due to aberrant RNA repression. If PUMs were underactive in tissues and did not properly regulate the mRNA encoding the DEK oncogene, which contains 3 PREs, the protein DEK may be overproduced. DEK is involved in DNA double strand break repair and is often overexpressed in many cancers. Overexpression of DEK inhibits cell death and functions in cellular survival (64). Thus, PUM depletion or loss of function may contribute to cancer pathology by causing upregulation of its target mRNAs.

Loss of PUM activity would also result in upregulation of its target mRNA, SMPDL3A, which has been linked to bladder tumorigenesis (61). The signal peptide SCUBE1, which contains a CUB-EGF domain, is expressed during early embryogenesis and is thought to be upregulated in heart disease and hypertension: recent studies employ SCUBE1 detection in diagnosis of Crimean-Congo hemorrhagic fever, viral illness that is potentially fatal and affects many regions of the world (60, 65). Intriguingly levels of SCUBE1, which is translocated to the surface of cells upon platelet stimulation, increase with worsening prognosis in CCHF, as the disease affects endothelial cells and

platelets(65). Perhaps a correlation between PUMs and SCUBE1 levels in viral infection exists; though this is purely speculative, it will be interesting to study direct regulation of SCUBE1 by PUMs in reporter assays.

Eight ncRNAs with no known function were identified in our deep sequencing analysis. Intriguingly, the ncRNAs LOC647979 and HCG11 have 15 and 16 PREs, respectively. This finding raises the question: what purpose might a ncRNA containing 15-16 PREs serve? Do these ncRNAs modulate PUM activity? Are they junk and might PUMs target them for removal? Both HCG11 and LOC647979 ncRNAs are intergenic non-protein coding RNAs, and are nearly identical in sequence but are on separate genes on different chromosomes. One possibility for their function follows closely with the ceRNA hypothesis, proposed recently to describe how long noncoding RNAs, mRNAs, and pseudogenes communicate with one another using miRNA elements (66). Perhaps noncoding RNAs such as LOC647979 and HCG11 serve to bind PUMs and compete for PUM targets as a means of regulating gene expression, much like miRNA sponges (67-69).

Together, our data implicate PUMs as important regulators of many biological pathways in humans, likely through specific interactions with PREs. We depleted PUM1 and PUM2 from human cells and discovered that they repress and activate nearly 500 RNAs that contain PRE sequences, including several noncoding RNAs. We propose that, in addition to their known repressive function, PUMs are capable activators of genes through a mechanism that has yet to be identified. Further, we suggest that they may directly regulate ncRNAs, which must be verified in future experiments. Together, our

data provide a high quality set of PUM regulated targets, and suggest biological pathways in which they may be involved.

References

1. Wickens, M., Bernstein, D. S., Kimble, J., and Parker, R. (2002) A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet.* **18**, 150–157
2. Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997) The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**, 1421–1433
3. Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickens, M. P. (1997) A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* **390**, 477–484
4. Wharton, R. P., Sonoda, J., Lee, T., Patterson, M., and Murata, Y. (1998) The Pumilio RNA-binding domain is also a translational regulator. *Molecular Cell* **1**, 863–872
5. Edwards, T. (2000) Crystallization and Characterization of Pumilio: A Novel RNA Binding Protein. *JOURNAL OF STRUCTURAL BIOLOGY* **132**, 251–254
6. Wang, X., Zamore, P. D., and Hall, T. M. (2001) Crystal structure of a Pumilio homology domain. *Molecular Cell* **7**, 855–865
7. Wang, X., McLachlan, J., Zamore, P. D., and Hall, T. M. T. (2002) Modular recognition of RNA by a human pumilio-homology domain. *Cell* **110**, 501–512
8. Wharton, R. P., and Struhl, G. (1991) RNA regulatory elements mediate control of Drosophila body pattern by the posterior morphogen nanos. *Cell* **67**, 955–967
9. Lin, H., and Spradling, A. C. (1997) A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development* **124**, 2463–2476
10. Parisi, M., and Lin, H. (1999) The Drosophila pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis. *Genetics* **153**, 235–250
11. Asaoka-Taguchi, M., Yamada, M., Nakamura, A., Hanyu, K., and Kobayashi, S. (1999) Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos. *Nature Publishing Group* **1**, 431–437
12. Barker, D. D., Wang, C., Moore, J., Dickinson, L. K., and Lehmann, R. (1992) Pumilio is essential for function but not for distribution of the Drosophila abdominal determinant Nanos. *Genes & Development* **6**, 2312–2326 [online] http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=1459455&retmode=ref&cmd=prlinks.
13. Crittenden, S. L., Rudel, D., Binder, J., Evans, T. C., and Kimble, J. (1997) Genes required for GLP-1 asymmetry in the early Caenorhabditis elegans embryo. *Dev Biol* **181**, 36–46
14. Dubnau, J., Chiang, A.-S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., Broger, C., and Tully, T. (2003) The staufen/pumilio pathway is involved in Drosophila long-term memory. *Current Biology* **13**, 286–296
15. Lehmann, R., and Nüsslein-Volhard, C. (1987) Involvement of the pumilio gene in the transport of an abdominal signal in the Drosophila embryo.
16. Mee, C. J., Pym, E. C. G., Moffat, K. G., and Baines, R. A. (2004) Regulation of neuronal excitability through pumilio-dependent control of a sodium channel gene.

Journal of Neuroscience **24**, 8695–8703
17. Menon, K. P., Sanyal, S., Habara, Y., Sanchez, R., Wharton, R. P., Ramaswami, M., and Zinn, K. (2004) The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E. *Neuron* **44**, 663–676
18. Nakahata, S., Kotani, T., Mita, K., Kawasaki, T., Katsu, Y., Nagahama, Y., and Yamashita, M. (2003) Involvement of Xenopus Pumilio in the translational regulation that is specific to cyclin B1 mRNA during oocyte maturation. *Mech. Dev.* **120**, 865–880
19. Rouhana, L., Wang, L., Buter, N., Kwak, J. E., Schiltz, C. A., Gonzalez, T., Kelley, A. E., Landry, C. F., and Wickens, M. (2005) Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA* **11**, 1117–1130
20. Schweers, B. A., Walters, K. J., and Stern, M. (2002) The Drosophila melanogaster translational repressor pumilio regulates neuronal excitability. *Genetics* **161**, 1177–1185
21. Vessey, J. P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M. A., and Macchi, P. (2006) Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. *Journal of Neuroscience* **26**, 6496–6508
22. Wreden, C., Verrotti, A. C., Schisa, J. A., Lieberfarb, M. E., and Strickland, S. (1997) Nanos and pumilio establish embryonic polarity in Drosophila by promoting posterior deadenylation of hunchback mRNA. *Development* **124**, 3015–3023
23. Chen, G., Li, W., Zhang, Q.-S., Regulski, M., Sinha, N., Barditch, J., Tully, T., Krainer, A. R., Zhang, M. Q., and Dubnau, J. (2008) Identification of Synaptic Targets of Drosophila Pumilio. *PLoS Computational Biology* **4**, e1000026
24. Xu, E. Y., Chang, R., Salmon, N. A., and Reijo Pera, R. A. (2007) 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
25. Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A. F., Elkon, R., and Agami, R. (2010) A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat Cell Biol* **12**, 1014–1020
26. Siemen, H., Colas, D., Heller, H. C., Brüstle, O., and Pera, R. A. R. (2011) Pumilio-2 function in the mouse nervous system. *PLoS ONE* **6**, e25932
27. Shigunov, P., Sotelo-Silveira, J., Kuligovski, C., de Aguiar, A. M., Rebelatto, C. K., Moutinho, J. A., Brofman, P. S., Krieger, M. A., Goldenberg, S., Munroe, D., Correa, A., and Dallagiovanna, B. (2012) PUMILIO-2 is involved in the positive regulation of cellular proliferation in human adipose-derived stem cells. *Stem Cells Dev.* **21**, 217–227
28. Chen, D., Zheng, W., Lin, A., Uyhazi, K., Zhao, H., and Lin, H. (2012) Pumilio 1 Suppresses Multiple Activators of p53 to Safeguard Spermatogenesis. *Curr Biol*
29. Miles, W. O., Tschöp, K., Herr, A., Ji, J.-Y., and Dyson, N. J. (2012) Pumilio facilitates miRNA regulation of the E2F3 oncogene. *Genes & Development* **26**, 356–368
30. Friend, K., Campbell, Z. T., Cooke, A., Kroll-Conner, P., Wickens, M. P., and

Kimble, J. (2012) A conserved PUF-Ago-eEF1A complex attenuates translation elongation. *Nat Struct Mol Biol* **19**, 176–183
31. Quenault, T., Lithgow, T., and Traven, A. (2011) PUF proteins: repression, activation and mRNA localization. *Trends Cell Biol.* **21**, 104–112
32. Goldstrohm, A. C., Seay, D. J., Hook, B. A., and Wickens, M. (2007) PUF protein-mediated deadenylation is catalyzed by Ccr4p. *J. Biol. Chem.* **282**, 109–114
33. Goldstrohm, A. C., Hook, B. A., Seay, D. J., and Wickens, M. (2006) PUF proteins bind Pop2p to regulate messenger RNAs. *Nat Struct Mol Biol* **13**, 533–539
34. Blewett, N. H., and Goldstrohm, A. C. (2012) A eukaryotic translation initiation factor 4E-binding protein promotes mRNA decapping and is required for PUF repression. *Molecular and Cellular Biology* **32**, 4181–4194
35. Chagnovich, D., and Lehmann, R. (2001) Poly(A)-independent regulation of maternal hunchback translation in the Drosophila embryo. *Proc Natl Acad Sci USA* **98**, 11359–11364
36. Gamberi, C., Peterson, D. S., He, L., and Gottlieb, E. (2002) An anterior function for the Drosophila posterior determinant Pumilio. *Development* **129**, 2699–2710
37. Suh, N., Crittenden, S. L., Goldstrohm, A., Hook, B., Thompson, B., Wickens, M., and Kimble, J. (2009) FBF and its dual control of gld-1 expression in the Caenorhabditis elegans germline. *Genetics* **181**, 1249–1260
38. Ahringer, J., and Kimble, J. (1991) Control of the sperm-oocyte switch in Caenorhabditis elegans hermaphrodites by the fem-3 3' untranslated region. *Nature* **349**, 346–348
39. Chritton, J. J., and Wickens, M. (2010) Translational repression by PUF proteins in vitro. *RNA* **16**, 1217–1225
40. Padmanabhan, K., and Richter, J. D. (2006) Regulated Pumilio-2 binding controls RINGO/Spy mRNA translation and CPEB activation. *Genes & Development* **20**, 199–209
41. Van Etten, J., Schagat, T. L., Hrit, J., Weidmann, C., Brumbaugh, J., Coon, J. J., and Goldstrohm, A. C. (2012) Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *Journal of Biological Chemistry*
42. Weidmann, C. A., and Goldstrohm, A. C. (2012) Drosophila Pumilio protein contains multiple autonomous repression domains that regulate mRNAs independently of Nanos and brain tumor. *Molecular and Cellular Biology* **32**, 527–540
43. Spassov, D. S., and Jurecic, R. (2002) Cloning and comparative sequence analysis of PUM1 and PUM2 genes, human members of the Pumilio family of RNA-binding proteins. *Gene* **299**, 195–204
44. Spassov, D. S., and Jurecic, R. (2003) The PUF family of RNA-binding proteins: does evolutionarily conserved structure equal conserved function? *IUBMB Life* **55**, 359–366
45. Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., and Gerber, A. P. (2008) Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System. *PLoS ONE* **3**, e3164

46. Morris, A. R., Mukherjee, N., and Keene, J. D. (2008) Ribonomic Analysis of Human Pum1 Reveals cis-trans Conservation across Species despite Evolution of Diverse mRNA Target Sets. *Molecular and Cellular Biology* **28**, 4093–4103
47. Lee, M.-H., Hook, B., Pan, G., Kershner, A. M., Merritt, C., Seydoux, G., Thomson, J. A., Wickens, M., and Kimble, J. (2007) Conserved regulation of MAP kinase expression by PUF RNA-binding proteins. *PLoS Genet* **3**, e233
48. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141
49. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Meth* **5**, 621–628
50. Livak, K. J., and Schmittgen, T. D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408
51. Schmittgen, T. D., and Livak, K. J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* **3**, 1101–1108
52. Thomas, P. D., Kejariwal, A., Guo, N., Mi, H., Campbell, M. J., Muruganujan, A., and Lazareva-Ulitsky, B. (2006) Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Research* **34**, W645–50
53. Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research* **13**, 2129–2141
54. Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* **37**, 1–13
55. Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57
56. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research* **40**, D109–14
57. Vessey, J. P., Schoderboeck, L., Gingl, E., Luzi, E., Riefler, J., Di Leva, F., Karra, D., Thomas, S., Kiebler, M. A., and Macchi, P. (2010) Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function. *Proceedings of the National Academy of Sciences* **107**, 3222–3227
58. Menon, K. P., Andrews, S., Murthy, M., Gavis, E. R., and Zinn, K. (2009) The translational repressors Nanos and Pumilio have divergent effects on presynaptic terminal growth and postsynaptic glutamate receptor subunit composition. *Journal of Neuroscience* **29**, 5558–5572
59. Salazar, A. M., Silverman, E. J., Menon, K. P., and Zinn, K. (2010) Regulation of synaptic Pumilio function by an aggregation-prone domain. *Journal of*

Neuroscience **30**, 515–522
60. Özkan, G., Ulusoy, S., Menteşe, A., Karahan, S. C., and Cansiz, M. (2013) New marker of platelet activation, SCUBE1, is elevated in hypertensive patients. *Am. J. Hypertens.* **26**, 748–753
61. Wright, K. O., Messing, E. M., and Reeder, J. E. (2002) Increased expression of the acid sphingomyelinase-like protein ASML3a in bladder tumors. *J. Urol.* **168**, 2645–2649
62. Bermudez, O., Jouandin, P., Rottier, J., Bourcier, C., Pagès, G., and Gimond, C. (2011) Post-transcriptional regulation of the DUSP6/MKP-3 phosphatase by MEK/ERK signaling and hypoxia. *J. Cell. Physiol.* **226**, 276–284
63. Piqué, M., López, J. M., Foissac, S., Guigó, R., and Méndez, R. (2008) A combinatorial code for CPE-mediated translational control. *Cell* **132**, 434–448
64. Kavanaugh, G. M., Wise-Draper, T. M., Morreale, R. J., Morrison, M. A., Gole, B., Schwemberger, S., Tichy, E. D., Lu, L., Babcock, G. F., Wells, J. M., Drissi, R., Bissler, J. J., Stambrook, P. J., Andreassen, P. R., Wiesmüller, L., and Wells, S. I. (2011) The human DEK oncogene regulates DNA damage response signaling and repair. *Nucleic Acids Research* **39**, 7465–7476
65. Menteşe, A., Yilmaz, G., Sümer, A., Arslan, M., Karahan, S. C., and Köksal, I. (2013) The diagnostic and prognostic significance of SCUBE1 levels in Crimean-Congo hemorrhagic fever. *Int. J. Infect. Dis.*
66. Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* **146**, 353–358
67. Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., and Pandolfi, P. P. (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038
68. Ebert, M. S., and Sharp, P. A. (2010) Emerging Roles for Natural MicroRNA Sponges. *Current Biology* **20**, R858–R861
69. Poliseno, L., Salmena, L., Riccardi, L., Fornari, A., Song, M. S., Hobbs, R. M., Sportoletti, P., Varmeh, S., Egia, A., Fedele, G., Rameh, L., Loda, M., and Pandolfi, P. P. (2010) Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation. *Science Signaling* **3**, ra29

Figures

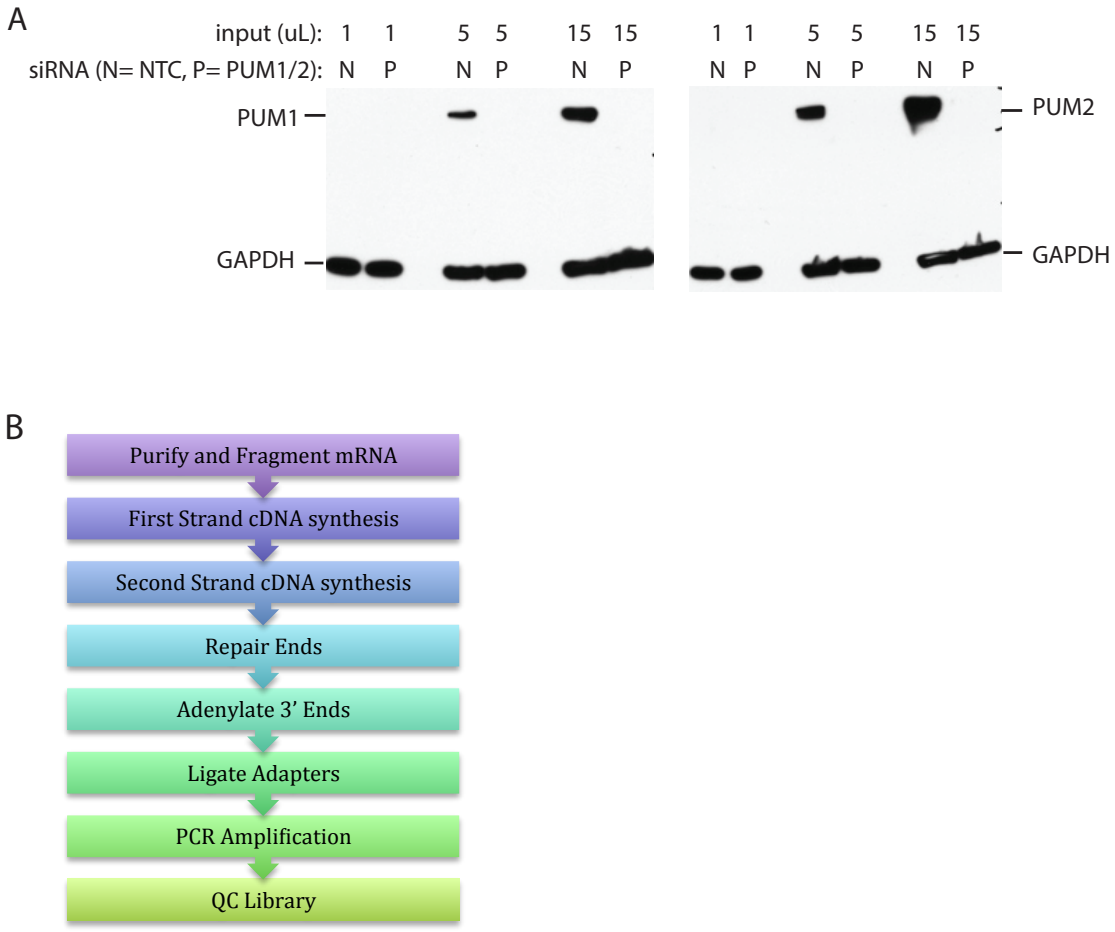


Figure 4.1. Library generation workflow and confirmation of PUM knockdown.

A. PUM knockdown was confirmed by Western blotting. Input lysates from either NTC or PUM1/2 siRNA treated cell populations were titrated from 1-15 µL and probed for either PUM1, PUM2, or GAPDH (loading control). B. Workflow for RNA Seq library generation. RNA was poly (A) selected and fragmented, followed by cDNA synthesis and adapter ligation. Lastly, cDNA fragments with barcode adapters were PCR amplified and subjected to analysis with a Bioanalyzer to confirm purity.

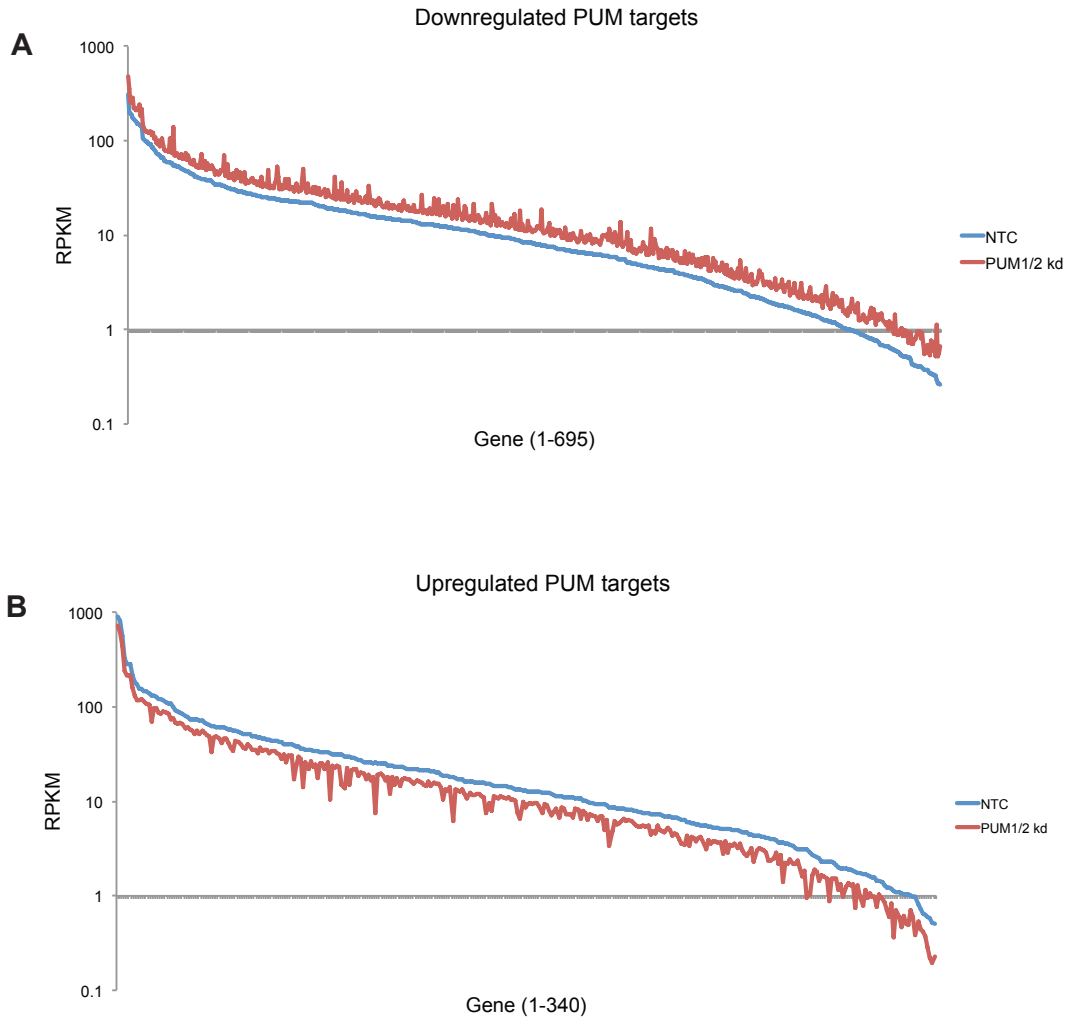


Figure 4.2. RPKM values of RNAs identified in RNA Seq span 7 orders of magnitude.

RPKM values of RNAs identified in NTC treated samples (blue line) and PUM1/2 siRNA treated samples (red line). A. Comparison of NTC to PUM1/2 kd RPKMs in 695 downregulated (repressed) RNAs. Note that RPKM values increased upon PUM knockdown. B. Comparison of NTC to PUM1/2 kd RPKM values for 340 upregulated (activated) RNAs. Note that, upon PUM knockdown, activated targets' RPKM values decreased.

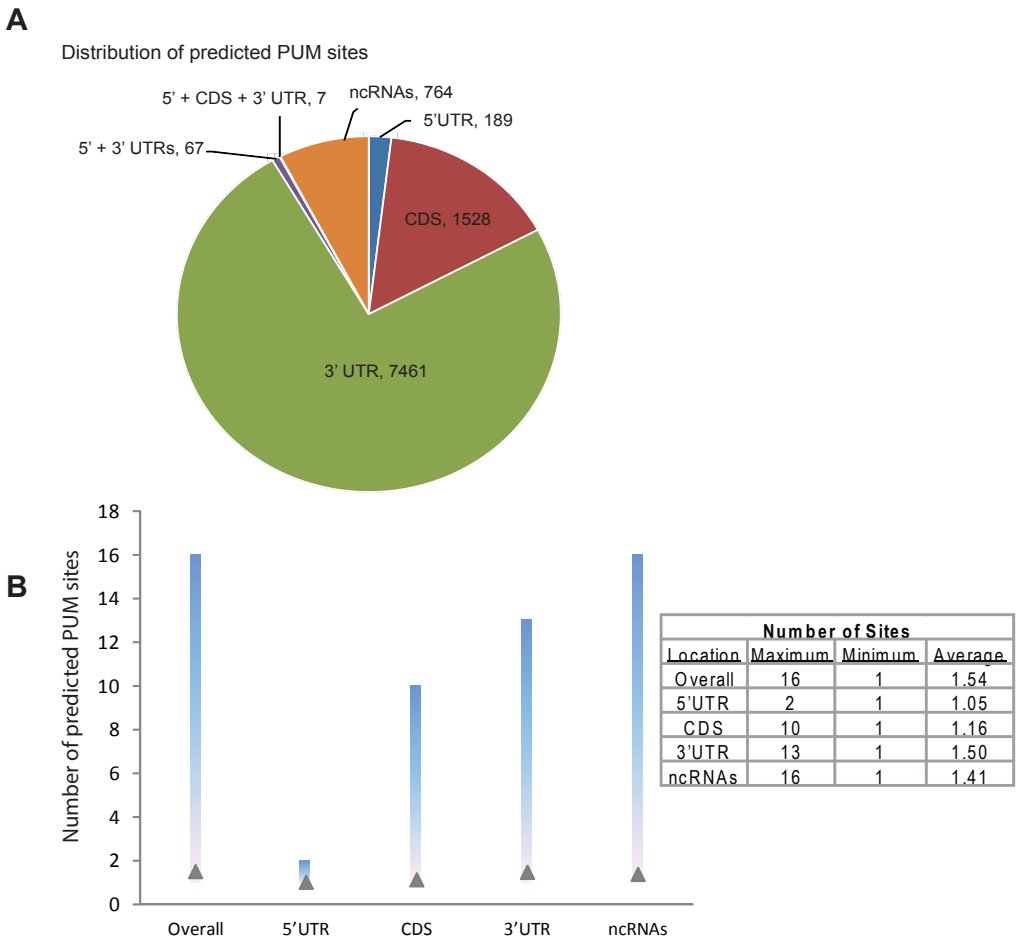


Figure 4.3. PRE predictions.

A. Distribution of the total number of predicted PREs in ncRNAs and mRNAs, organized either by their location. B. Bar graph representing the number of PREs per transcript, divided by location. Grey triangles represent the average number of PREs per region. Blue gradient bars represent the range in total number of PREs per transcript.

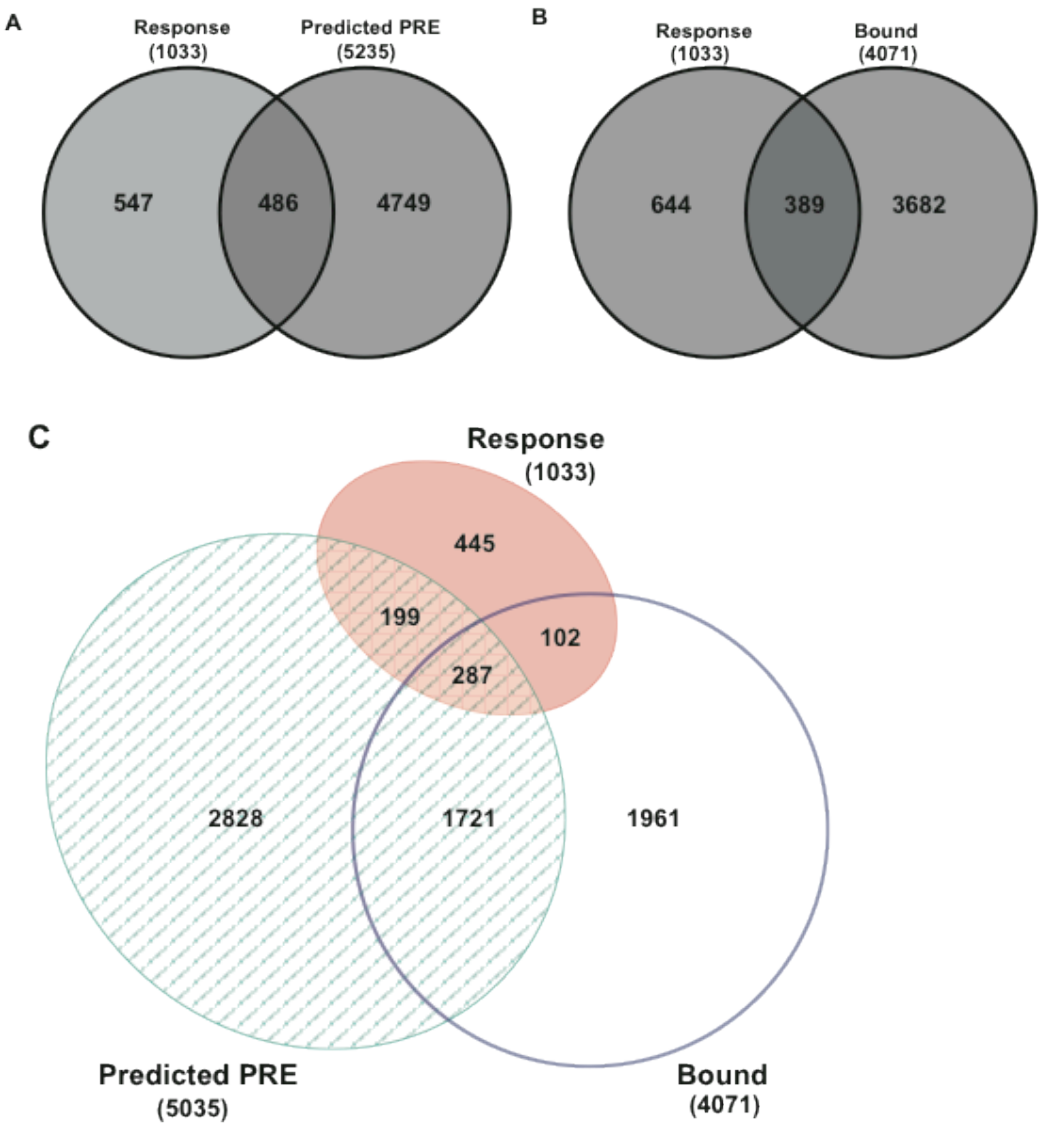


Figure 4.4. Cross reference of predicted PREs, Bound RNAs, and Response RNAs.

A. 1033 differentially expressed genes, "Response," from our dataset were compared with 5235 RNAs predicted to have PREs, "Predicted PRE." 486 contain a PRE and are differentially regulated in our dataset. B. 1033 "Response" genes were compared with 4071 unique RNAs identified in RIP-Chip and PAR-CLIP studies, called the "Bound" dataset. 389 RNAs were differentially regulated and identified in the "Bound" dataset. C. Cross comparison of all three datasets, "Response," "Predicted PRE," and "Bound." Of those RNAs, 287 were differentially regulated in our dataset, contain at least one PRE, and were previously identified in IP experiments.

Tables

Total number of reads per sample	26 million x 6 samples
Total number of genes identified	
Control	18685
PUM k/d	18582
Range of RPKM values	
Control	0.000924-1681
PUM k/d	0.001311-1788
Total # differentially expressed genes	1035
Total UP	695
Total DOWN	340
Range fold change	3.81-0.287
Total # genes changed >1.5 fold	387
UP	271
DOWN	116

Table 4.1. RNA-Seq parameters and experimental statistics.

More than 26 million reads were obtained in each of 6 samples (3 control, 3 PUM1/2 knockdown) from 18685 and 18582 total genes, respectively. The range of RPKM values spanned 7 orders of magnitude in this analysis for both control and PUM k/d samples.

Biological Process	# genes in ref list (20000)	# genes in set	# genes expected	Over/under?	P-value
cellular process	6072	403	290.55	+	1.75E-12
cell communication	4224	288	202.12	+	7.88E-09
signal transduction	4019	274	192.31	+	3.10E-08
developmental process	2840	205	135.89	+	1.72E-07
system development	1911	149	91.44	+	4.97E-07
nervous system development	1146	98	54.84	+	5.75E-06
ectoderm development	1347	109	64.45	+	1.49E-05
cell-cell adhesion	724	69	34.64	+	1.58E-05
cell adhesion	1301	106	62.25	+	1.64E-05
cell-matrix adhesion	207	27	9.9	+	7.99E-04
mesoderm development	1347	101	64.45	+	1.23E-03
protein modification process	1330	96	63.64	+	8.78E-03
protein transport	1542	107	73.78	+	1.49E-02
intracellular protein transport	1542	107	73.78	+	1.49E-02
cell surface receptor linked signal transduction	2049	135	98.04	+	1.74E-02
transport	2679	169	128.19	+	1.94E-02
Molecular Function	**# genes in ref list (20000)**	**# genes in set**	**# genes expected**	**Over/under?**	**P-value**
protein binding	3073	212	147.04	+	2.87E-06
receptor binding	1172	93	56.08	+	2.78E-04
metallopeptidase activity	233	28	11.15	+	2.05E-03
receptor activity	1852	128	88.62	+	2.96E-03
enzyme regulator activity	1216	89	58.19	+	9.04E-03
phosphatase activity	239	25	11.44	+	4.73E-02
Cellular Component	**# genes in ref list (20000)**	**# genes in set**	**# genes expected**	**Over/under?**	**P-value**
extracellular matrix	558	61	26.7	+	1.73E-07
extracellular region	601	63	28.76	+	4.47E-07

Table 4.2. PANTHER gene ontology analysis of 1033 differentially regulated PUM targets.

1033 differentially expressed genes from our dataset were entered into PANTHER db and analyzed for statistical overrepresentation among three different GO categories: Biological Process (blue), Molecular Function (green), and Cellular Component (purple) (52, 53). Shown are GO terms that are statistically represented among PUM targets with p-values ≤0.05. The second column lists the total number of classified genes in the reference list, containing 20,000 genes. The third column lists the total number of genes in our dataset that match the corresponding GO analysis term listed in the first column. The fourth column lists the total number of predicted genes in each GO term category from the reference list. The fifth column indicates whether the GO term is overrepresented or underrepresented in our dataset. The last column lists p-values for each GO term represented in our dataset. Note that only GO terms with p-values ≤0.05 are listed here.

KEGG Pathway Term	Count	%	PValue	Fold Enrichment
Axon guidance	21	2.05	1.29E-04	2.59
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	15	1.46	2.16E-04	3.14
Regulation of actin cytoskeleton	28	2.73	3.58E-04	2.07
Dilated cardiomyopathy	16	1.56	5.19E-04	2.76
Hypertrophic cardiomyopathy (HCM)	15	1.46	7.14E-04	2.80
Aldosterone-regulated sodium reabsorption	9	0.88	3.39E-03	3.49
ECM-receptor interaction	12	1.17	1.51E-02	2.27
MAPK signaling pathway	27	2.63	1.57E-02	1.61
Gap junction	12	1.17	2.24E-02	2.14
p53 signaling pathway	10	0.97	2.50E-02	2.34
Endocytosis	19	1.85	3.87E-02	1.64
Chemokine signaling pathway	19	1.85	4.44E-02	1.61
TGF-beta signaling pathway	11	1.07	4.44E-02	2.01
Melanogenesis	12	1.17	4.46E-02	1.93

Table 4.3. KEGG pathway analysis.

1033 differentially expressed genes from our dataset were entered into the DAVID functional analysis tool and analysed for statistical overrepresentation using KEGG pathway analysis (54-56). Shown are KEGG terms that are overrepresented in our dataset (first column). The second column lists the total number of genes in our dataset that correspond to the given KEGG term. The fourth column lists the p-value: shown here are KEGG terms with p-values ≤ 0.05. Fold enrichment of each KEGG term in our dataset is listed in the last column.

Biological Process	# genes in ref list (20000	# genes in set	# genes expect	Over/under	P-value
cellular process	6072	295	198.55	+	2.69E-13
cell communication	4224	224	138.12	+	1.06E-12
signal transduction	4019	216	131.42	+	1.08E-12
cell adhesion	1301	98	42.54	+	3.38E-12
cell-cell adhesion	724	65	23.67	+	9.12E-11
system development	1911	114	62.49	+	5.52E-08
nervous system development	1146	79	37.47	+	9.63E-08
developmental process	2840	148	92.87	+	8.50E-07
ectoderm development	1347	85	44.05	+	1.21E-06
cell-matrix adhesion	207	23	6.77	+	1.09E-04
cell surface receptor linked signal transduction	2049	105	67	+	5.14E-04
neurological system process	1917	99	62.69	+	7.52E-04
mesoderm development	1347	75	44.05	+	1.05E-03
system process	2159	107	70.6	+	1.74E-03
transport	2679	123	87.6	+	1.11E-02
heart development	286	23	9.35	+	1.73E-02
cell-cell signaling	1259	66	41.17	+	2.26E-02
visual perception	416	29	13.6	+	2.68E-02
transmembrane receptor protein tyrosine kinase signaling pathway	295	23	9.65	+	2.69E-02
protein transport	1542	77	50.42	+	2.81E-02
intracellular protein transport	1542	77	50.42	+	2.81E-02
protein modification process	1330	68	43.49	+	3.62E-02
Molecular Function	**# genes in ref list (20000**	**# genes in set**	**# genes expect**	**Over/under**	**P-value**
receptor activity	1852	113	60.56	+	1.58E-08
protein binding	3073	154	100.49	+	4.36E-06
receptor binding	1172	73	38.32	+	2.25E-05
metallopeptidase activity	233	23	7.62	+	6.62E-04
enzyme regulator activity	1216	65	39.76	+	1.29E-02
guanyl-nucleotide exchange factor activity	162	16	5.3	+	1.81E-02
small GTPase regulator activity	490	32	16.02	+	3.46E-02
Cellular Component	**# genes in ref list (20000**	**# genes in set**	**# genes expect**	**Over/under**	**P-value**
extracellular matrix	558	57	18.25	+	3.62E-12
extracellular region	601	59	19.65	+	6.73E-12

Table 4.4. GO analysis of repressed PUM targets.

PANTHER was used to calculate statistical enrichment of genes in specific GO categories in repressed PUM targets. 406 upregulated RNAs from our dataset were entered into PANTHER db and analysed for statistical overrepresentation (52, 53). Shown are GO terms that are stastically overrepresented in the RNA-Seq dataset with p-values <0.05. The second column lists the total number of classified genes in the reference list, containing 20,000 genes. The third column lists the total number of genes in our dataset that match the corresponding GO analysis term listed in the first column. The fourth column lists the total number of predicted genes in each GO term category from the reference list. The fifth column indicates whether the GO term is overrepresented or underrepresented in our dataset. The last column lists p-values for each GO term represented in our dataset. Note that only GO terms with p-values ≤0.05 are listed here.

GENE NAME	FC RNASEQ	FC qPCR	SEM qPCR
ANO4	2.295	2.1	0.2
DEK	1.629	1.4	0.1
DUSP6	0.478	0.8	0.0
ETV4	0.297	0.3	0.0
FMR1	1.267	1.4	0.2
FZD8	2.215	1.5	0.2
L1CAM	1.622	1.5	0.1
NOVA2	1.978	2.1	0.1
RET	1.944	1.9	0.1
SCUBE1	2.379	2.9	0.2
SMPDL3A	2.35	2.3	0.2

Table 4.5. qRT-PCR validation of RNA Seq targets.

Relative fold change of each validated PUM target RNA in our RNA Seq dataset (FC RNASEQ) (listed in column 2). The third column lists relative fold change calculated by qRT-PCR of each target (FC qPCR). mRNA derived from total RNA isolated form cells transfected with PUM1/2 siRNA versus NTC siRNA. qPCR assays were done in duplicate from triplicate RNA samples to control for pipetting error, transfection variability, and sample-to-sample variation. The last column gives standard error of the mean values for qPCR studies (SEM qPCR).

Gene ID	Proposed/known biological role
ANO4	Anoctamin 4, calcium activated chloride channel
DEK	Oncogene, contains SAP domain, binds DNA to induce positive supercoils, splice site selection, upreg assoc. with disease
DUSP6	Member of dual specificity protein phosphatase family- inactivates ERK2, possible role in several cancers, possible oncogene
ETV4	Role in branching morphogenesis in developing kidney
FMR1	Fragile X mental retardation, premature ovarian failure
FZD8	Receptors coupled to beta-catenin signaling, upregulated in some cancers
L1CAM	Corpus callosum, partial agenesis of, CRASH syndrome, Hirschprung disease, MASA syndrome
NOVA2	Autoantigen in paraneoplastic opsoclonus myoclonus ataxia associated with breast cancer, fallopian cancer
RET	Proto-oncogene involved in central hypoventilation syndrome, multiple endocrine neoplasisas, Hirschsprung disease.
SCUBE1	Cell surface glycoprotein, may play role in vascular biology
SMPDL3A	Acid sphingomyelinase-like phosphodiesterase 3a, upregulated in bladder cancer

Table 4.6. Known and proposed biological roles of validated PUM targets.

We chose targets for qPCR validation that are putative oncogenes and neural regulators, as well as those RNAs that were changed most in RNA-Seq studies, including SCUBE1 and SMPDL3A. We determined "proposed/known biological role" from annotations in the NCBI GENE database.

Gene ID	PUF sites	Motif sequences	3' UTR	5' UTR	CDS
ANO4	3	tgtacata, tgtacata, tgtaaata	3	0	0
DEK	3	tgtaaata, tgtatata, tgtatata	2	0	1
DUSP6	2	tgtaaata, tgtaaata	2	0	0
ETV4	1	tgtacata	1	0	0
FMR1	2	tgtagata, tgtacata	2	0	0
FZD8	1	tgtatata	1	0	0
L1CAM	2	tgtaaata, tgtacata	2	0	0
NOVA2	1	tgtatata	1	0	0
RET	1	tgtaaata	1	0	0
SCUBE1	1	tgtaaata	1	0	0
SMPDL3A	2	tgtagata, tgtaaata	2	0	0

Table 4.7. Number and location of PREs in validated PUM targets.

All PUM targets chosen for validation contain at least one PRE. In the second column, we list the total number of PREs in each target gene. In the third column, we list the UGUANAUA sequence(s) found in each PUM target. In the last three columns, the distribution of predicted PREs is listed. Note that the majority of predicted PREs in PUM targets are enriched in the 3' UTR, with the exception of DEK, which has an additional PRE in its coding region (CDS).

CHAPTER 5

Perspectives and Outlook

Summary

Several decades of work on the PUF family of RNA binding regulators have provided tremendous insight into mechanisms of posttranscriptional gene regulation in eukaryotes. PUFs in model organisms are repressors of mRNAs that bind the PRE and repress mRNA targets via several proposed mechanisms, and are generally correlated with mRNA decay. The high degree of homology between PUMs and other PUF family members' RNA binding domains motivates a widespread acceptance of a role of human PUMs as repressors. Yet, limited data exists that verifies direct regulation PUM targets in cells. Given the tremendous potential for PUM regulation in posttranscriptional regulation, as evidenced by the many hundreds of putative PUM target genes identified in RIP-ChIP and PAR-CLIP binding studies in mammals and in model organisms, we found it crucial to not only investigate the modes of PUM regulation in human cells, but to identify regulated RNA targets of both human PUMs.

Our studies demonstrate that mammalian PUMs may regulate many hundreds of mRNAs and give insight into mechanisms of PUM mediated repression in human cells.

In Chapter 2 we describe the development of luciferase-based assays that are well suited to quantitatively measure 3' UTR directed regulation in cells (1). Together these analyses offer a streamlined approach to studying posttranscriptional regulatory mechanisms, which can be readily applied to studying other RNA regulatory systems. We provide a detailed guide to the construction of luciferase reporters that have RNA binding target sequences in their 3'UTRs, as well as to detect regulated luciferase reporter protein and mRNA levels. We provide insight into the regulatory mechanisms of both canonical human PUMs in Chapter 3: PUM1 and PUM2 can repress the same reporter mRNA, and each possesses potent intrinsic repression activity mediated by specific interactions with the PRE

PUFs in model organisms act in concert with corepressor proteins to carry out RNA repression. PUFs in yeast cause shortening of the poly (A) tail and reduction in mRNA levels: they are thought to enhance deadenylation and decay via a conserved interaction with the POP2 subunit of the CCR4-POP2 deadenylase complex (2-4). In chapter 3 we demonstrate that an interaction between human PUMs and POP2 orthologs, CNOT7 and CNOT8, is vital for PUMs to carry out efficient repression of a target mRNA, which, in addition to prior *in vitro* evidence, provides further evidence that the POP2-PUF interaction is conserved in humans (3, 4). We observe a correlation between repression of luciferase protein and mRNA levels, indicating that human PUMs enhance mRNA decay, presumably through recruitment of the CCR4-NOT deadenylase complex via an interaction with CNOT7 or CNOT8. Our results extend evidence in model organisms to humans that brings to light a conserved model for PUF repression; in which PUFs recruit the POP2 subunit of the CCR4-NOT deadenylase complex to shorten the

poly (A) tail, enhancing mRNA decay and promoting translational repression. We also demonstrate that PUMs repress via a deadenylation-independent mechanism, which results in repression of a luciferase reporter that has a histone stem loop in place of a poly (A) tail, and find that these mRNAs are modestly stabilized when examined by qRT-PCR.

In addition to deciphering mechanisms of human PUM repression, we identified over 1000 differentially regulated RNAs that respond to changes to a reduction in PUM expression by RNAi-mediated silencing in HEK293 cells. 486 differentially regulated genes boast at least 1 PRE: intriguingly, about 15% of those targets were decreased upon PUM knockdown, meaning they are potentially activated by PUMs. This suggests a potential role for PUMs as activators, for which there is limited precedence in the literature (4, 5). We then demonstrate validation of several PRE- containing differentially expressed RNA targets by qRT-PCR analysis.

Collectively, our findings shed light on the potential for broad-reaching regulatory control by PUM1 and PUM2 proteins in many biological processes and demonstrate their direct regulation of several RNAs in human cells.

Perspectives on PUM regulatory mechanisms and combinatorial control with corepressors

Deadenylation-independent repression

We demonstrated that PUMs repress via a conserved interaction with subunits of the CCR4-NOT deadenylase complex in Chapter 3. Interestingly, we discovered that PUMs repress even in the absence of a poly (A) tail, which suggests that PUMs repress using multiple mechanisms. Several studies implicate PUMs as direct regulators of

translation (1, 6, 7). Replacement of the poly (A) tail with a histone stem loop on the RnLUC 3xPRE reporter mRNA allowed the process of deadenylation to be bypassed and presumably, be decapped and degraded (2, 8). Importantly, the RnLUC 3xPRE HSL reporter was repressed by human PUMs, albeit to a lesser degree than the RnLUC 3xPRE p(A) reporter: therefore, PUMs cause repression in a manner independent of the poly(A) tail and consequently, of deadenylation. At first glance, the most likely scenario for repression involves an interaction between PUM and decapping factors, including Dcp1 and Dcp2, to the 5' end of the mRNA to remove the cap, which would be in alignment with normal histone metabolism. The decapped mRNA would then be subjected to rapid decay by exoribonuclease activity. Alternatively, PUMs may direct translational repression of the RnLUC 3xPRE HSL through another undetermined mechanism.

Our evidence indicates that deadenylation-independent regulation of the RnLUC PRE HSL reporter occurs at the level of translation and not by decapping and subsequent mRNA decay, as RnLUC PRE HSL mRNAs are stabilized despite luciferase protein repression. One intriguing possibility is that human PUMs interact directly with the cap or with 4EBPs to interfere with translation initiation. 4EBPs are characterized by a YXXXLφ motif, where X is any residue and φ is hydrophobic, which they share with eIF4G: this motif is essential for the interaction between 4G or 4EBPs with eIF4E (3, 4, 9). 4EBPs typically compete with eIF4G for binding to eIF4E: when bound to eIF4E, 4EBPs block translation by preventing 4G binding, formation of the 4F complex, and stimulation of translation initiation (4, 10). An interaction between an eIF4E orthologs and a PUF protein promotes binding of the 4EBP to eIF4E (5, 11). Perhaps an interaction

between a human PUM and a 4EBP strengthens the interaction between the 4EBP and eIF4E, thereby reducing translation of the HSL reporter.

An alternative mechanism for PUM regulation, though not likely in our RnLUC HSL reporter assays, is that 5' decapping and degradation of a PUM RNA target occurs: there is precedent for 5' decapping and decay initiated by PUF protein binding to a PRE in yeast. Yeast Puf5p requires the 4E binding protein (4EBP) EAP1 to promote decapping and mRNA degradation of the HO mRNA, rather than to inhibit translation initiation directly (12). Further work will be necessary to understand if 4EBPs are involved in human PUM-mediated repression of mRNA targets. It is also possible that PUMs compete with the 5' cap to prevent eIF4E binding: Xenopus PUM2 outcompetes eIF4E to bind to the 5' cap and subsequent formation of the 4F complex in order to reduce translation (13).

Convergence of miRNA and PUM regulatory pathways

PUM and miRNA repression pathways involve recruitment of many of the same corepressors: for example, human PUMs and miRISC bind subunits of the CCR4-NOT deadenylase complex (2, 14). Recent work postulates that PUMs cooperate with the miRNA regulatory pathway (15-18). *In vitro* experiments in a recent study demonstrated that a complex consisting of PUM2, Ago, and eIF1A formed to block translation elongation independently of miRNA association (16). It is important to note that this study, while carried out with human PUM2, only looked at regulation conferred by the RBD. Further, this analysis did not address PUM1.

We know that both PUMs are highly active and that they are widely expressed in tissues. Furthermore, there exists evidence that the *Drosophila* PUM confers the majority

of its repression through repressive 3 domains in its N terminus (19). When tethered, human PUM1 and PUM2 N termini repress a luciferase reporter in *Drosophila* cells, providing further evidence that the AGO-eEF1A-PUM translational inhibitory complex, while a tantalizing explanation for PUM mediated repression, is not the only mode by which human PUMs repress target mRNAs (19). Our laboratory demonstrated recently that neither human nor *Drosophila* PUMs require the interaction with the AGO-eEF1A inhibitory complex to carry out repression and that PUMs repress independently (2) (Weidmann et al., submitted). Furthermore, we showed that the PUM RBD in *Drosophila* enhances deadenylation through an interaction with PABP, which is entirely independent of the effects of binding either AGO or eEF1A (Weidmann et al., submitted). While a plausible mechanism in *in vitro* biochemical assays, it does not contribute, or contributes very weakly, to PUM repression in human and *Drosophila* cells (Weidmann et al., submitted).

Another study describing the interplay between miRNAs and human PUMs postulates that PUM1 binds the mRNA encoding the PRE-containing tumor suppressor p27 to induce a structural change, leaving miR221/222 sites accessible for binding and repression (15). p27 downregulation is required for entry into the cell cycle from quiescent cells and thus, the repression mediated by PUM1 and miRNAs does point to a role for human PUM1 in cell proliferation, which is consistent with PUM function in other studies and PUF function in model organisms (17, 20).

Taken together, these mechanisms are intriguing: miRNAs and PUMs recruit many of the same decay factors, like the CCR4-NOT complex, so it is likely the pathways converge and possibly cooperate. However, cooperative binding between

miRNAs and PUMs cannot be the sole mechanism of human PUM repression. Both PUMs are highly active repressors of a *Renilla* luciferase reporter in human cells, which I discussed in Chapter 3. Furthermore, the base pairing between the PRE and miR221/222 is unique: no other miRNA seed sites are complementary to the PRE. PUFs in model organisms are repressors of many mRNAs and do so in the absence of the miRNA pathways: PUFs in yeast, for example, are potent repressors do not compete with miRNAs as there is no equivalent miRNA – induced silencing pathway in the organism. It is likely, however, that since 8 nucleotide PRE sites are predicted to be abundant in human mRNAs, which we discussed in Chapter 4, and miRNA sites are reported to be enriched surrounding PREs, cooperativity between PUMs and miRNAs is probable under certain conditions (18). Further work should focus on interrogating a direct cooperative interaction between pathways in cells: a simple way to test involvement of the miRNA pathway in PUM repression is to place known miRNA binding sites in proximity of the PRE in the RnLUC 3xPRE reporter. It will thusly be possible to measure synergism between miRNA binding sites and the PRE. From that data, we will gain a better understanding of cooperativity between miRNAs and PUMs, which will provide a strong foundation on which the nuances of PUM-miRNA cooperativity in posttranscriptional regulation can be studied.

RNA target activation by PUMs

We discovered that 340 genes were decreased by PUM knockdown in RNA Seq experiments; of those, 80 contain predicted PREs that are located in the 5' and 3' UTRs and in the coding region, and 35 were also in the Bound data set. This is indicative of direct activation of those targets by PUMs. We confirmed PUM dependent regulation of

two “activated” PUM targets, DUSP6 and ETV4, by qRT-PCR. There are two possible mechanisms to explain this activation: first, PUMs act directly on the mRNA and cause its activation by an unknown mechanism, likely by recruitment of another protein cofactor that activates translation. Second, loss of PUMs may upregulate a gene that typically represses the “activated” mRNA, thereby causing its downregulation upon PUM knockdown. The latter indirect model is less likely for mRNAs that have PREs and are bound by PUMs. The indirect model is a more likely explanation for the mRNAs that are downregulated and do not have PREs.

Several published observations support the potential ability of PUFs to activate mRNAs. FBF in *C. elegans* is proposed to cause activation of translation of the *egl-4* mRNA via binding to its PRE in response to environmental stimuli (21). Furthermore, PUF9 in *T. brucei* stabilizes several mRNAs during S phase (22). One proposed mechanism of PUF activation is that *C. elegans* FBF binds the poly (A) polymerase GLD-2 to promote polyadenylation and activation of translation of the gld-1 mRNA (5). Perhaps PUM1 or PUM2 bind a GLD-2 type poly (A) polymerase, which will thereby extend the poly (A) tail to promote translation. In another example, CPEB in *Xenopus* assembles with different RBP complexes to either activate or repress translation. In *Xenopus*, upon insertion of a PRE near a cytoplasmic polyadenylation element, translation is increased 2-fold: it is thought that this activation is caused by stabilization of the CPE binding protein CPEB on the RNA (23). Given this evidence in *Xenopus*, one possibility is that PREs in proximity to CPEs in the UTR work in a combinatorial fashion to recruit proteins that activate translation. A simpler model is that PUMs could bind and

displace another dominant repressor, resulting in "activation," which is in actuality, de-repression.

In order to test direct activation by PUMs, we can create luciferase reporters that encode the entire ORF and 3'UTR of a potential target gene: for example, DUSP6. Then, we can mutate its PREs to determine if the specific interaction between PUMs and PRE is required for gene activation. It is also imperative that direct binding is demonstrated; this can be done employing gel shift assays with recombinant human PUM RBD and *in vitro* transcribed RNA encoding the gene of interest. It is important that this tests whether PUMs *can* cause activation of these genes using PRE sequences: whether they do so in the native setting cannot be elucidated using these assays. Demonstrating indirect regulatory control of activated genes may be more difficult to parse out if specific repressors are unknown: by taking a candidate gene approach, one might identify the intermediary protein(s) between activated RNA and PUF protein. Activity of those proteins can be modulated by creating inactive mutants or by depleting them with RNAi in concert with PUM knockdown to measure effects on target RNA activation. It may be that different RNA regulators exert combinatorial control to activate gene expression or downregulate gene expression in response to other *cis* elements in different locations in the mRNA.

Widespread regulation of biological processes by PUMs

Identification of RNA targets provided our laboratory and the research community with a tremendous amount of information regarding potential PUM targets in human cells. Very few direct targets of PUF regulation in model organisms and in humans have been validated: we confirmed regulation of 11 genes by qRT-PCR and

recognize that many more have yet to be validated. qRT-PCR provided a validation of the data obtained in RNA Seq data: analysis of mRNA decay and protein repression is necessary to confirm PUM target regulation.

PUMs regulate hundreds of mRNAs including noncoding RNAs

We demonstrated regulation of over 1000 RNAs, including messenger RNAs and ncRNAs, upon knockdown of PUMs in human cells: the RNAs we identified are involved in several biological pathways, including p53 signaling, axon guidance, and development. If PUM control fails or becomes hyperactive, one can imagine that regulation of many pathways would be greatly perturbed, leading to many diseases, including cancer. It is possible that PUMs control regulatory pathways by controlling individual components of many pathways, or by regulating several components within the same pathway. From our gene ontology data, it appears that PUMs repress individual RNAs across many biological regulatory pathways; though in some pathways, PUMs regulate many RNAs. For example, in the p53 signaling pathway, PUMs regulate six different mRNAs genes in the pathway, including p21 (CDKN1A), p53R2 (RRM2B), PUMA (BBC3), Cyclin G (CCNG2), CDC2 (CDK1), and CASP3, all of which contain at least one PRE.

We discovered 8 ncRNAs that were regulated by PUMs in our RNA Seq analysis: 3 of those contain PREs including NEAT1, LOC647979, and HCG11. None of the identified ncRNAs have a known function. Intriguingly, LOC647979 and HCG11 contain 15 and 16 PREs, respectively. The observation that 2 ncRNAs containing 15 and 16 PREs is puzzling: do these ncRNAs regulate PUM function in some way? Can they act as PUM sponges, in a manner akin to miRNA sponges? Do PUMs target their PREs in an

effort to clear them away because they are junk sequences? Many of these questions should be addressed in future experiments. One could easily adapt our luciferase based reporter assays to measure repression of (or by) these intriguing ncRNAs.

We validated regulation of several genes identified in disease pathways including FMR1, DEK, RET, and ETV4, by qRT-PCR. Only a few genes have been identified as direct RNA targets of human PUMs, though evidence for functional regulation of those genes is limited, and very little is known about the biological implications of regulation of those RNA targets by PUM proteins. It has long been thought that PUMs bind very specialized subsets of genes in specific categories related to function, for example, development. Our results and those of others shed light on the possibility of extensive regulatory potential by PUMs in humans and other mammals (18, 20, 24, 25). We showed that PUMs appear to regulate multiple components of specific diverse pathways. Future work will examine the roles and functions of PUM targets in humans, as well as the distinct mechanisms by which they are regulated by PUM proteins and their corepressors. Given their high degree of sequence similarity and overlapping functions, it is likely that PUM1 and PUM2 are essential in mammals for proper development and function, though this remains to be seen. It is important to note that in reporter assays, PUM1 and PUM2 bind the same sequence and repress the same target, and both must be knocked down to lose repression in HEK293 cells. Altered expression levels of PUMs in different cells and tissues may distinguish specific roles for PUM1 and PUM2. It will be important to attempt to generate a PUM1/2 double knockout mouse in the future to understand whether PUMs are essential for development. If viable, a double knockout mouse would provide tremendous insight into effects on gross morphology and

development. However, in the likely case that a double knockout mutant is not viable, tissue specific depletion of PUMs could be utilized.

Outlook and future directions

Through our work, we provide a demonstration of direct regulation by PUM1 and PUM2 in human cells: importantly, we established that each PUM is a repressor, and identified corepressor proteins that bind directly to PUMs to enhance mRNA decay. We also opened new avenues of inquiry into PUM function in humans: PUMs regulate many hundreds of RNAs including mRNAs and ncRNAs, which play roles in a broad range of biological processes.

Many questions remain unanswered in terms of understanding the mechanisms of PUM regulation. We showed that one facet of PUM repression involves a conserved interaction with the CCR4-NOT deadenylase complex, but another pathway is deadenylation-independent. It is important to continue to explore and characterize PUMs' interactions with members of the deadenylase complex: CNOT7 and CNOT8 are active deadenylases, as are the CCR4 orthologs in humans. This differs from yeast in that POP2 is thought to be a bridging interaction between CCR4 and the PUF protein (3, 4). Furthermore, since humans have two orthologs each of CCR4 and POP2, do PUMs preferentially interact with specific deadenylase subunits? We have evidence that PUM1 and PUM2 bind all four subunits, though we have not yet measured their binding affinities. Deadenylase subunits form heterodimeric complexes, by which CNOT6 or CNOT6L bind to either CNOT7 or CNOT8: we do not yet know if PUMs prefer a particular configuration of POP2/CCR4 subunits within the larger deadenylase complex. We also do not fully understand which subunit carries out deadenylation: all four

subunits are thought to be catalytically active, which is different from the PUF-POP2 deadenylation mechanism in yeast which points to CCR4 being the primary catalytic subunit for PUF-mediated deadenylation (3). To date, we have been unable to demonstrate a direct shortening of the poly (A) tail in human cells, as the reporter mRNA we used was very low in abundance, and thus efforts to characterize its poly (A) tail length by RNase cleavage and Northern blot analysis were unsuccessful. Now that we have identified and validated RNA targets in HEK293 cells, an alternative approach to using a reporter mRNA will involve measuring the effects of PUMs on poly (A) tail length of endogenous target RNA.

In addition to a conserved deadenylation dependent mechanism, PUMs use a deadenylation-independent mechanism, which is less clearly understood. Histone mRNAs are typically degraded by 5' decapping and decay, we initially suspected that, if PUMs repressed the RnLUC 3xPRE HSL reporter at all, we would see a commensurate reduction in mRNA levels of the reporter. We did not observe a reduction of reporter mRNA, and instead measured its stabilization despite a marked reduction in luciferase reporter protein. In this case, it is proposed that PUMs inhibit translation and do not affect mRNA levels. Several studies implicate PUFs in direct translational control and poly (A) independent mRNA repression, though a particular mechanism for such control remains to be elucidated (6, 7, 19).

One likely scenario is that PUMs bind different corepressor proteins and complexes. Early in this project we pursued identification of members of the deadenylase complex as PUM corepressors because deadenylase complex subunits were identified in purified PUM complexes by mass spectrometry: our efforts yielded the finding that

PUMs interact with deadenylase subunits of the CCR4-NOT complex to cause a reduction of reporter protein and mRNA levels. In light of this, it is likely that many more corepressors bind to PUMs such as miRNAs, Nanos, and translational regulators like DAZL, DAZ, and BOL, for which there is evidence in the literature in mammals (15-18, 26-28). Decapping and decay regulators and 4EBPs are possible corepressors in PUM regulation as well, though there is limited evidence for these interactions in model organisms and none in humans. A straightforward way to test involvement of candidate corepressor proteins is to systematically knock them down by RNAi in human cells and use the methods described in Chapter 2 to measure regulation of a reporter gene such as RnLUC 3xPRE. Coimmunoprecipitations to measure direct binding interactions as well as mutational analyses of corepressors proteins will provide useful experimental systems in which to measure their interactions with PUMs.

We know that both fly PUM and human PUM RBDs confer very little repression in comparison with the full length or N-terminal region of PUMs alone (19). The fly PUM N terminus confers the majority of PUM repression in DMEL-2 cells and is composed of three modular and independent regulatory domains (19). How these domains cause repression is currently unknown, but studies to elucidate the mechanisms of fly and human PUM N termini are underway in our laboratory and will be useful in gaining a holistic understanding of PUM regulatory mechanisms. Furthermore, our laboratory recently demonstrated that the PUM RBD, which seems to repress by causing deadenylation, depends on an interaction with PABP to confer regulatory activity (Weidmann et al., submitted). It will prove useful in humans to interrogate the PABP – poly (A) tail- PUM connection to determine if the PUM RBD in humans represses using

a similar mechanism and to see if deadenylation-dependent and independent mechanisms in PUMs are attributable to separable activities of protein domains.

We identified PREs in the 3' UTR of all genes we validated by qRT-PCR, though the mRNA encoding the DEK oncogene contains a PRE in its coding region. No one has demonstrated regulation of a RNA target through a direct interaction between a PUF protein and PRE in a location other than the 3'UTR. Future experiments should focus on understanding whether the location of the PRE affects the ability of PUMs to cause repression. Our task from this point forward is to demonstrate functional regulation of activated and repressed PUM targets in cells. I constructed several clones that take advantage of a luciferase cell-based assay format: the ORF and 3' UTR of a target RNA was cloned into the pF5A mammalian expression vector and a NanoLuc luciferase ORF was inserted upstream of the target stop codon. I systematically mutated each PRE in the gene. Using these clones, which have been created for several putative PUM biological targets, we will measure direct regulation by human PUMs as a function of luciferase activity. Furthermore, creation of stable PUM1/2 knockdowns may be most useful in terms of studying the effects of PUM depletion on target RNAs. Similarly, regulation of ncRNAs identified in our RNA Seq analysis may shed light on the function of those ncRNAs, which is previously unknown, as well as on the potential for a new class of PUM targets. Alternatively, we may learn that PUMs are regulated by noncoding RNAs under certain conditions. PUM interactions through PREs located in ncRNAs highlights the urgency for the continued study of ncRNA functions and their biological significance.

PUMs confer widespread regulation of target genes and cause repression in a number of ways. Future work will focus on defining poly (A) independent PUM

regulatory mechanisms and the involved corepressors, as well as the characterization of PUM mediated repression of newly identified RNA targets. Together, these avenues of inquiry will yield important information regarding mechanisms of posttranscriptional regulatory control of a broad range of biological pathways, and shed light on normal physiology and development, and how misregulation of posttranscriptional regulation conferred by PUMs contributes to human disease. Our insights into PUM direct targets, combined with evidence from other laboratories, are, colloquially, the tip of the iceberg in terms of providing insight into the pervasiveness of PUM regulation in human biology.

References

1. Van Etten, J., Schagat, T. L., and Goldstrohm, A. C. (2013) A guide to design and optimization of reporter assays for 3' untranslated region mediated regulation of mammalian messenger RNAs. *Methods*
2. Van Etten, J., Schagat, T. L., Hrit, J., Weidmann, C., Brumbaugh, J., Coon, J. J., and Goldstrohm, A. C. (2012) Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *Journal of Biological Chemistry*
3. Goldstrohm, A. C., Seay, D. J., Hook, B. A., and Wickens, M. (2007) PUF protein-mediated deadenylation is catalyzed by Ccr4p. *J. Biol. Chem.* **282**, 109–114
4. Goldstrohm, A. C., Hook, B. A., Seay, D. J., and Wickens, M. (2006) PUF proteins bind Pop2p to regulate messenger RNAs. *Nat Struct Mol Biol* **13**, 533–539
5. Suh, N., Crittenden, S. L., Goldstrohm, A., Hook, B., Thompson, B., Wickens, M., and Kimble, J. (2009) FBF and its dual control of gld-1 expression in the Caenorhabditis elegans germline. *Genetics* **181**, 1249–1260
6. Chagnovich, D., and Lehmann, R. (2001) Poly(A)-independent regulation of maternal hunchback translation in the Drosophila embryo. *Proc Natl Acad Sci USA* **98**, 11359–11364
7. Chritton, J. J., and Wickens, M. (2010) Translational repression by PUF proteins in vitro. *RNA* **16**, 1217–1225
8. Marzluff, W. F., Wagner, E. J., and Duronio, R. J. (2008) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature Reviews Genetics* **9**, 843–854
9. Cridge, A. G., Castelli, L. M., Smirnova, J. B., Selley, J. N., Rowe, W., Hubbard, S. J., McCarthy, J. E. G., Ashe, M. P., Grant, C. M., and Pavitt, G. D. (2010) Identifying eIF4E-binding protein translationally-controlled transcripts reveals links to mRNAs bound by specific PUF proteins. *Nucleic Acids Research* **38**, 8039–8050
10. Lasko, P. (2003) Gene regulation at the RNA layer: RNA binding proteins in intercellular signaling networks. *Sci. STKE* **2003**, RE6
11. Kong, J., and Lasko, P. (2012) Translational control in cellular and developmental processes. *Nature Reviews Genetics* **13**, 383–394
12. Blewett, N. H., and Goldstrohm, A. C. (2012) A eukaryotic translation initiation factor 4E-binding protein promotes mRNA decapping and is required for PUF repression. *Molecular and Cellular Biology* **32**, 4181–4194
13. Cao, Q., Padmanabhan, K., and Richter, J. D. (2010) Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition. *RNA* **16**, 221–227
14. Fabian, M. R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J. T., Svitkin, Y. V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J. A., Chen, C.-Y. A., Shyu, A.-B., Yates, J. R., Hannon, G. J., Filipowicz, W., Duchaine, T. F., and Sonenberg, N. (2009) Mammalian miRNA RISC recruits CAF1 and PABP to

affect PABP-dependent deadenylation. *Molecular Cell* **35**, 868–880
15. Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A. F., Elkon, R., and Agami, R. (2010) A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat Cell Biol* **12**, 1014–1020
16. Friend, K., Campbell, Z. T., Cooke, A., Kroll-Conner, P., Wickens, M. P., and Kimble, J. (2012) A conserved PUF-Ago-eEF1A complex attenuates translation elongation. *Nat Struct Mol Biol* **19**, 176–183
17. Miles, W. O., Tschöp, K., Herr, A., Ji, J.-Y., and Dyson, N. J. (2012) Pumilio facilitates miRNA regulation of the E2F3 oncogene. *Genes & Development* **26**, 356–368
18. Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., and Gerber, A. P. (2008) Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System. *PLoS ONE* **3**, e3164
19. Weidmann, C. A., and Goldstrohm, A. C. (2012) Drosophila Pumilio protein contains multiple autonomous repression domains that regulate mRNAs independently of Nanos and brain tumor. *Molecular and Cellular Biology* **32**, 527–540
20. Chen, D., Zheng, W., Lin, A., Uyhazi, K., Zhao, H., and Lin, H. (2012) Pumilio 1 Suppresses Multiple Activators of p53 to Safeguard Spermatogenesis. *Curr Biol*
21. Kaye, J. A., Rose, N. C., Goldsworthy, B., Goga, A., and L'Etoile, N. D. (2009) A 3'UTR pumilio-binding element directs translational activation in olfactory sensory neurons. *Neuron* **61**, 57–70
22. Archer, S. K., Luu, V.-D., de Queiroz, R. A., Brems, S., and Clayton, C. (2009) Trypanosoma brucei PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. *PLoS Pathog.* **5**, e1000565
23. Piqué, M., López, J. M., Foissac, S., Guigó, R., and Méndez, R. (2008) A combinatorial code for CPE-mediated translational control. *Cell* **132**, 434–448
24. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141
25. Morris, A. R., Mukherjee, N., and Keene, J. D. (2008) Ribonomic Analysis of Human Pum1 Reveals cis-trans Conservation across Species despite Evolution of Diverse mRNA Target Sets. *Molecular and Cellular Biology* **28**, 4093–4103
26. Fox, M., Urano, J., and Reijo Pera, R. A. (2005) Identification and characterization of RNA sequences to which human PUMILIO-2 (PUM2) and deleted in Azoospermia-like (DAZL) bind. *Genomics* **85**, 92–105
27. Moore, F. L., Jaruzelska, J., Fox, M. S., Urano, J., Firpo, M. T., Turek, P. J., Dorfman, D. M., and Pera, R. A. R. (2003) Human Pumilio-2 is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-like proteins. *Proc Natl Acad Sci USA* **100**, 538–543
28. Urano, J., Fox, M. S., and Reijo Pera, R. A. (2005) Interaction of the conserved meiotic regulators, BOULE (BOL) and PUMILIO-2 (PUM2). *Mol. Reprod. Dev.* **71**, 290–298

APPENDIX

Supplemental Information for RNA-Seq Analysis

Table S1

GENE	RPKM (Control)	RPKM (PUM1/2 kd)	Fold change	Diff exp >1.5-fold?
PUM2	25.989	7.4502	0.286667449	YES
ETV4	3.08002	0.950251	0.308520872	YES
SPRY4	1.14	0.368495	0.323240528	YES
AKR7A2	32.4012	10.5002	0.324068343	YES
TAGLN	18.0101	6.28508	0.348973647	YES
NPPB	2.83305	1.04636	0.369342196	YES
PKDREJ	0.526295	0.195009	0.370531988	YES
ETV5	2.30748	0.872925	0.378301919	YES
PUM1	36.4423	14.0578	0.385755866	YES
HMX3	3.6385	1.41581	0.389118101	YES
TRIL	0.576203	0.224433	0.389503987	YES
ANXA2P3	0.975128	0.382048	0.391792042	YES
EDEM2	8.72348	3.41955	0.391993055	YES
CPA4	1.81808	0.740677	0.407394949	YES
FBXO36	1.06666	0.4683	0.439032608	YES
ANKRD54	39.051	17.3484	0.444248866	YES
LOC643650	0.501647	0.22974	0.457971587	YES
ELAVL1	30.0266	13.9118	0.463316511	YES
RSAD2	1.70738	0.79476	0.465486102	YES
VPS52	0.589519	0.283107	0.480234182	YES
TFDP3	1.0376	0.499759	0.481647639	YES
C17orf82	1.02729	0.49816	0.484927172	YES
CHPT1	31.1883	15.1517	0.485811989	YES
SLC25A15	15.204	7.43301	0.488886001	YES
DKFZP686I15217	1.20543	0.590892	0.490192393	YES
RASSF3	13.3576	6.63614	0.496807893	YES
ETV1	1.93204	0.972053	0.503122634	YES
TTC3P1	23.9511	12.0571	0.503404843	YES
TTC3	29.5493	15.1453	0.512544439	YES
ZCCHC3	33.812	17.4599	0.516381063	YES
DUSP6	3.08956	1.61869	0.523921292	YES
INSIG1	131.917	69.1701	0.524343811	YES
SUMF2	62.9593	33.1322	0.526248456	YES
PKI55	4.36594	2.30284	0.527455395	YES
MTNR1A	1.45027	0.769935	0.530890415	YES
KRT80	9.26404	4.93654	0.532871655	YES
SERTAD2	14.8359	7.94807	0.535730746	YES
IFIT3	6.35219	3.44035	0.54160126	YES

ABHD16B	5.12081	2.79036	0.544905516	YES
SLC25A34	1.0857	0.598647	0.551393819	YES
SGCB	13.4939	7.51438	0.556873289	YES
C20orf177	8.69808	4.86385	0.559187494	YES
CTXN1	18.8998	10.6067	0.561208202	YES
REEP6	6.44213	3.62487	0.56268211	YES
FBXO27	2.00753	1.14222	0.568970595	YES
GNRHR2	5.95493	3.39258	0.569709749	YES
PDSS2	5.49316	3.12992	0.569785179	YES
CCL2	5.06051	2.90032	0.573127887	YES
ST3GAL3	4.65369	2.67202	0.57417324	YES
PAMR1	1.21775	0.702847	0.577171113	YES
SKAP1	1.64831	0.952544	0.577891678	YES
SYT12	1.72696	1.00889	0.584199893	YES
MPZL3	3.98694	2.34413	0.587951069	YES
C3orf34	3.72287	2.19009	0.588278412	YES
OAS2	1.03959	0.612577	0.589250507	YES
LOC100129961	2.46859	1.45628	0.589924814	YES
LOC729013	4.05352	2.39492	0.590824272	YES
DCPS	12.0051	7.09995	0.591410187	YES
DRD1	0.624081	0.372137	0.59629494	YES
TFDP2	7.37994	4.40573	0.596986826	YES
MX1	2.75153	1.64288	0.59707911	YES
LOC100129726	3.08076	1.83949	0.597091113	YES
NCRNA00173	4.67817	2.80688	0.599995672	YES
CREBL2	12.4518	7.47144	0.600026865	YES
ZNF853	11.3275	6.82871	0.602844575	YES
CCDC146	0.993198	0.599957	0.604065543	YES
FAM172A	56.6166	34.266	0.60522983	YES
DQX1	0.721548	0.437087	0.605763266	YES
DOHH	10.9571	6.64389	0.606356011	YES
PCNX	10.6281	6.49348	0.610971505	YES
ACYP2	3.54819	2.17298	0.612418598	YES
C19orf66	1.58135	0.980583	0.620091492	YES
SH3RF1	6.47879	4.0229	0.620934071	YES
CCND1	38.1383	23.7836	0.623615633	YES
IFI6	35.1568	21.9515	0.624388989	YES
LCOR	7.3101	4.57394	0.625700863	YES
PRELID2	3.96683	2.48661	0.626851667	YES
LRRC23	3.54828	2.22965	0.628376007	YES
FBXO9	23.1386	14.5677	0.629583659	YES
LHPP	6.82504	4.29763	0.62968622	YES
IFIT2	3.11947	1.96537	0.630033305	YES
FOS	1.55645	0.983684	0.632003302	YES
TEAD2	23.5422	14.9807	0.636334391	YES
TRAPPC9	5.54526	3.53792	0.638008257	YES
PARK2	0.651414	0.417558	0.641003438	YES
CAPN10	15.1851	9.755	0.642405879	YES
HOXA10	40.5569	26.0838	0.643139677	YES
APH1B	4.48193	2.88343	0.643345665	YES
MMP24	3.40647	2.19258	0.643649863	YES
TXNDC12	56.9345	36.9596	0.649160445	YES
IFI35	4.68195	3.04107	0.64953042	YES
ZNF599	2.31339	1.50267	0.64955023	YES
DOK3	12.2072	7.92953	0.649576344	YES
NFIC	25.6107	16.6515	0.65017861	YES
C11orf75	9.90017	6.44143	0.650638005	YES
IL6R	4.94823	3.22142	0.651025519	YES
PROM2	1.095	0.71393	0.65199328	YES

IFIT1	11.0639	7.21618	0.652226968	YES
KILLIN	1.30703	0.852809	0.65248064	YES
PITPNC1	9.39018	6.13552	0.653397572	YES
INTS5	11.0179	7.21637	0.654966489	YES
TMCC2	5.77851	3.79674	0.657043594	YES
ATL3	11.1197	7.30966	0.657363382	YES
QRFPR	1.93176	1.27025	0.65756162	YES
LRRC20	5.43459	3.57885	0.658531794	YES
MSRB3	11.7024	7.7117	0.65898293	YES
PRORSD1P	1.619	1.06886	0.660195404	YES
SP140L	2.28515	1.50908	0.66038534	YES
VPS26A	26.0107	17.1866	0.660748431	YES
LOC728392	16.3156	10.7927	0.661498138	YES
CDH13	0.801462	0.531196	0.662783228	YES
DHRS1	6.72319	4.45624	0.662815387	YES
TGFBI	1.40745	0.934916	0.664261406	YES
TSPYL1	42.138	28.0255	0.665089316	YES
FAM101B	33.0099	21.9643	0.665386269	YES
HS1BP3	8.4819	5.65405	0.666601745	YES
RASEF	25.0612	16.719	0.66712592	NO
C15orf41	7.3255	4.89625	0.668384415	NO
LOC257396	1.90353	1.2787	0.671750782	NO
CCT6B	2.25283	1.51572	0.672809042	NO
ACTBL2	6.7996	4.60356	0.677034849	NO
LOC729678	1.96109	1.32777	0.677055028	NO
PEX6	12.8813	8.73474	0.678092973	NO
CACNG4	3.17942	2.15647	0.67825891	NO
TMEM117	6.99284	4.74604	0.678700979	NO
DAB1	1.69921	1.15516	0.679825791	NO
C2CD2	4.97979	3.39834	0.682426701	NO
SH3GL1	47.515	32.4514	0.682971366	NO
ZNF71	7.2345	4.94399	0.68338998	NO
CFDP1	60.8163	41.626	0.68445519	NO
NXPH3	1.02048	0.699378	0.685340571	NO
UBE2L3	34.0035	23.3605	0.687001427	NO
GDI2	108.192	74.3291	0.687014284	NO
MAML1	21.0559	14.5016	0.688717862	NO
DDOST	172.765	118.99	0.68873839	NO
ASTN1	5.05615	3.49122	0.690490275	NO
LBH	21.1819	14.6617	0.692182801	NO
PET112L	8.41948	5.82892	0.692313315	NO
PANK1	9.10322	6.31104	0.693275653	NO
ISG15	27.1766	18.8485	0.693555865	NO
SOCS5	7.07006	4.90509	0.693783772	NO
PPIP5K1	4.36462	3.02989	0.694192652	NO
MYL12A	121.369	84.4685	0.695967094	NO
TECR	51.8277	36.1916	0.698305867	NO
ZNF365	2.31353	1.61839	0.699534444	NO
MRPL1	7.78732	5.44853	0.699667314	NO
BRD2	3.45206	2.41537	0.699689623	NO
CECR2	7.87411	5.51776	0.700747696	NO
PHF15	4.05801	2.84437	0.700925978	NO
SOX3	5.20449	3.65362	0.702013169	NO
BIRC5	73.4246	51.552	0.702108549	NO
PER2	1.18554	0.832629	0.702318333	NO
CRLF3	15.7641	11.0857	0.703223408	NO
SNX8	7.72	5.42962	0.703318953	NO
SAAL1	21.3235	14.998	0.703356004	NO
CBWD1	22.0773	15.5612	0.704850404	NO

MNAT1	25.586	18.0667	0.706119371	NO
CPEB1	1.92013	1.35611	0.706262793	NO
LOC147727	12.619	8.92592	0.707341592	NO
PCDHB2	2.54299	1.80197	0.708604238	NO
C21orf119	18.4108	13.0549	0.709091644	NO
SLC9A3	2.31848	1.64517	0.709589713	NO
BAZ2A	31.5117	22.4158	0.711348778	NO
STXBP6	5.24842	3.7397	0.712538069	NO
ZFHX4	5.88501	4.19694	0.713157186	NO
RAB43	1.85935	1.32634	0.713331209	NO
ABHD14B	6.88762	4.91785	0.714013373	NO
TOP1	52.0962	37.2358	0.714750188	NO
NRG2	2.6206	1.87354	0.714927078	NO
CALCOCO1	33.9941	24.3101	0.715127308	NO
SLC46A3	5.80666	4.1531	0.715230914	NO
H2AFY	182.715	130.801	0.715876685	NO
TTC33	6.00999	4.30472	0.716260853	NO
SRD5A3	5.21996	3.74014	0.716507147	NO
CAP2	15.8192	11.3369	0.716650195	NO
SQRDL	4.97221	3.56389	0.716762468	NO
FHL1	122.496	87.8151	0.716881715	NO
MAP2K6	5.43931	3.90004	0.717009928	NO
DUSP19	1.43708	1.03059	0.717142637	NO
C13orf1	5.13176	3.68757	0.718578662	NO
COMMD8	8.53449	6.14594	0.720129863	NO
SFN	7.4979	5.39947	0.720130362	NO
WDR1	72.5845	52.3253	0.720886981	NO
CTDNEP1	44.8475	32.3687	0.721750947	NO
SKA1	17.3774	12.5558	0.722535311	NO
SGK196	5.68903	4.11815	0.723875252	NO
ZNF658	1.77134	1.28294	0.724274757	NO
UBASH3B	4.23548	3.06868	0.724518785	NO
CLASP1	16.1039	11.6678	0.724530335	NO
NEFH	4.40025	3.18913	0.724759377	NO
PEPD	46.8801	34.0022	0.725301632	NO
FAM195B	48.4997	35.1855	0.725478618	NO
TSTD1	49.0053	35.5639	0.725715003	NO
CCNT1	18.6821	13.5794	0.726866842	NO
ZNF768	57.4892	41.8693	0.728299621	NO
KRT18	131.865	96.368	0.730807325	NO
ELFN2	3.92596	2.86977	0.730972481	NO
LRRC57	7.53693	5.51243	0.731389084	NO
ARHGAP31	2.09397	1.53472	0.732922724	NO
DIAPH3	12.828	9.40404	0.733085309	NO
SCFD2	4.88975	3.58513	0.733193042	NO
C11orf24	22.1268	16.2366	0.733799588	NO
STAMBPL1	7.93635	5.82629	0.734128237	NO
RPH3AL	4.34885	3.19293	0.734202026	NO
HSPA2	15.9917	11.7475	0.734599592	NO
LOC401022	9.34575	6.86807	0.734887848	NO
RYK	16.334	12.0196	0.735864479	NO
CYR61	35.2461	25.9518	0.736302753	NO
SRP14	148.341	109.252	0.736488549	NO
FAM64A	28.9161	21.2976	0.736531432	NO
SCN1B	9.27167	6.82936	0.736583508	NO
ZW10	14.6492	10.7959	0.736967038	NO
HOXA11-AS1	9.4885	7.00165	0.737909094	NO
JARID2	21.324	15.7612	0.739128957	NO
IGF2BP2	20.5351	15.1801	0.739229892	NO

AOC2	4.26438	3.15391	0.739593782	NO
PSMD8	213.8	158.369	0.740734789	NO
ITGB3	3.69308	2.73564	0.740747625	NO
KRT17	288.454	213.78	0.741124591	NO
ZNF217	10.0982	7.48724	0.741443158	NO
FOXN3	11.4222	8.47179	0.741697597	NO
GOLGA7	31.3575	23.2682	0.742030298	NO
NT5C2	12.2932	9.12592	0.742355944	NO
STOX2	3.14984	2.3385	0.742418209	NO
CORO2B	7.41926	5.50845	0.742452688	NO
NDUFS2	51.3509	38.1305	0.742546357	NO
CLIP3	15.9486	11.8501	0.743013846	NO
PPP2R2D	27.0822	20.1231	0.743040112	NO
SERTAD4	20.4827	15.2633	0.745179556	NO
RPTOR	14.3953	10.7298	0.745371726	NO
TNIK	12.4082	9.24884	0.745382059	NO
ZNF618	12.7416	9.49788	0.745424426	NO
CYB5R3	88.7258	66.1425	0.745470413	NO
MSN	91.7235	68.4077	0.745802739	NO
PAF1	140.14	104.524	0.745854953	NO
CCDC50	11.1275	8.30095	0.745983176	NO
CUX1	23.1332	17.2574	0.746004377	NO
UBE2QL1	5.11783	3.81863	0.74614297	NO
LPAR1	10.801	8.06325	0.746528891	NO
LOC283070	8.30606	6.2016	0.746636012	NO
C5orf55	8.24971	6.16014	0.746710022	NO
PTMS	325.675	243.236	0.746867901	NO
FAM116A	12.7026	9.50792	0.748501442	NO
CASP6	9.7308	7.28546	0.748701215	NO
ZFYVE9	19.0161	14.2384	0.74875415	NO
TMEM129	25.7447	19.2766	0.748760378	NO
APPBP2	17.4154	13.0403	0.74878062	NO
GNAI2	61.0029	45.7024	0.749182965	NO
EGR1	9.50354	7.12684	0.749914487	NO
WBP2	74.4181	55.8159	0.750031452	NO
HDAC8	15.9026	11.9276	0.750042369	NO
STAT1	43.5343	32.7133	0.75143749	NO
SLC2A3	26.0248	19.5583	0.751524999	NO
NFE2L1	79.499	59.8002	0.752213446	NO
TRIM25	25.0829	18.8839	0.752861297	NO
MGAT5B	5.62915	4.24036	0.753284631	NO
RNF26	31.3624	23.6437	0.753884283	NO
STYXL1	13.5067	10.1893	0.754394467	NO
MRPL40	66.898	50.4784	0.754557631	NO
FAF1	33.6678	25.4054	0.754590582	NO
SPOP	24.0706	18.1861	0.75553317	NO
BCKDHA	42.2144	31.9134	0.75598473	NO
SCD5	23.2565	17.5906	0.756372596	NO
AGPAT3	33.0493	25.0177	0.756981528	NO
IRF9	16.1765	12.2476	0.757125834	NO
USP7	35.0094	26.5119	0.757280141	NO
STXBP1	17.8846	13.5454	0.757376729	NO
WWP2	13.9995	10.6037	0.757435529	NO
BGLAP,PMF1,PMF1-BGLAP	40.1308	30.4171	0.757949167	NO
SIVA1	62.8628	47.6474	0.757958624	NO
HERPUD2	12.9728	9.84924	0.75922163	NO
KRT8	286.96	218.042	0.759834435	NO
PTPN18	13.7005	10.4169	0.760332835	NO
DAG1	23.7148	18.0336	0.760435611	NO

PKD2	12.7949	9.73027	0.760483578	NO
SMARCE1	73.2634	55.719	0.760528912	NO
ZNF777	14.6566	11.1714	0.762210295	NO
HADH	28.75	21.9183	0.762376735	NO
NEFM	45.7308	34.8658	0.762413727	NO
MGAT5	12.6314	9.63068	0.76243698	NO
GNG12	31.5275	24.0409	0.762535812	NO
HMGA2	40.4182	30.8234	0.762612456	NO
TPM3	53.4881	40.0162	0.763007953	NO
TKT	128.237	97.8695	0.763189379	NO
NT5DC2	98.8474	75.4402	0.763197844	NO
PCBD1	60.4779	46.1739	0.763483032	NO
FAM70A	21.4587	16.3957	0.764059555	NO
VAPB	14.4161	11.0411	0.765889943	NO
PAK1	25.2077	19.3117	0.766102854	NO
C6orf106	45.3315	34.7409	0.766374254	NO
RAB6A	29.8483	22.8856	0.766730247	NO
DCP1A	6.89152	5.2853	0.766928506	NO
TULP4	7.29704	5.59835	0.767208708	NO
FOXO4	14.2316	10.92	0.767309222	NO
NCS1	21.86	16.7819	0.767698641	NO
MESDC1	20.7611	15.9637	0.768926177	NO
NUP93	43.9756	33.8424	0.769571886	NO
NEAT1	149.058	114.718	0.769617228	NO
KIAA1191	64.3084	49.5624	0.770698773	NO
MRFAP1L1	36.6082	28.2385	0.771371634	NO
UBE2M	44.297	34.1739	0.771472694	NO
PSMB6	154.793	119.432	0.771559863	NO
SSH2	11.3234	8.7371	0.771599439	NO
FARP1	22.1551	17.1111	0.772332507	NO
MED22	22.6148	17.4792	0.772907141	NO
PPP2R2A	22.1917	17.1606	0.773288145	NO
PPIL1	51.7686	40.0519	0.773670409	NO
EXOC4	28.4735	22.0347	0.773866171	NO
GNG4	8.05969	6.24037	0.774269115	NO
MPDU1	40.8017	31.595	0.774355525	NO
KHSRP	116.593	90.2937	0.774431747	NO
BICC1	14.7525	11.4255	0.774476302	NO
PKM2	801.999	621.495	0.774932739	NO
AHCYL2	10.2158	7.92341	0.775605533	NO
PRMT6	21.125	16.3873	0.775728117	NO
GINS2	56.5719	43.9352	0.776627124	NO
LRRC8A	25.1773	19.5919	0.778158529	NO
PCGF3	14.5572	11.3305	0.778345176	NO
DHCR24	108.833	84.7726	0.778925363	NO
SLAIN1	17.1388	13.3538	0.779157559	NO
FNBP1	18.2578	14.256	0.780815718	NO
NUP188	55.2041	43.119	0.781083127	NO
PARP14	8.29984	6.48484	0.781321381	NO
SMARCD1	46.9407	36.676	0.781325172	NO
TUBA1B	573.956	448.633	0.781650184	NO
FLOT2	84.5473	66.2383	0.783446682	NO
NUAK1	10.8229	8.48144	0.783656325	NO
RTN3	59.7573	46.8328	0.783716078	NO
RBFOX2	38.0633	29.8368	0.783873631	NO
DCTN1	68.1023	53.3931	0.784014369	NO
SSU72	74.4234	58.3516	0.784050236	NO
XPR1	8.10493	6.36609	0.785459068	NO
GARS	61.4128	48.251	0.785682864	NO

UBE2MP1	86.7567	68.1902	0.785993889	NO
C18orf10	71.5982	56.3016	0.786355725	NO
MKI67	111.739	87.92	0.786836615	NO
PFN1	916.025	722.299	0.788514943	NO
NOL11	32.8284	26.0018	0.79205197	NO
RAB1B	77.7463	61.6319	0.792730286	NO
CCT4	154.037	122.426	0.794782493	NO
CKAP5	70.6424	88.2448	1.249176211	NO
PRPS2	72.8874	91.6413	1.257300063	NO
GLG1	27.8911	35.0847	1.257915489	NO
PLEC	11.1275	14.0449	1.262172416	NO
SDHA	97.7836	123.702	1.265059288	NO
LOC642852	11.8651	15.0154	1.265508327	NO
TTF2	30.4995	38.6572	1.267472108	NO
LRP11	28.9115	36.6518	1.26772164	NO
MARK4	17.2863	21.9564	1.270158026	NO
KDM6A	22.5181	28.6036	1.270250472	NO
HIATL1	34.4226	43.7692	1.271525152	NO
EIF4B	52.4846	66.8031	1.272814345	NO
LOC647979	59.4641	75.7592	1.274033312	NO
DENND4C	12.9397	16.5007	1.275196876	NO
ADCY7	11.3949	14.5489	1.276796852	NO
MBD2	40.6258	51.8834	1.277103987	NO
PTPRA,VPS16	74.2084	94.8548	1.278221625	NO
SLC12A2	14.3343	18.3367	1.279218759	NO
DPY19L3	25.2803	32.3432	1.279385467	NO
KIF2A	34.1468	43.6981	1.279710965	NO
SACS	14.3967	18.4376	1.28068618	NO
ATF3	49.6225	63.565	1.280972052	NO
H1F0	97.6946	125.168	1.281216249	NO
ARID2	13.6675	17.5116	1.281257101	NO
HMGCR	38.6245	49.5154	1.281967777	NO
ODZ3	14.793	18.9686	1.282269935	NO
SEMA3C	13.0278	16.7107	1.282692185	NO
EPT1	15.6244	20.0463	1.28301052	NO
IRS2	6.61538	8.49105	1.283530875	NO
AIP	44.7051	57.4092	1.284177832	NO
FMR1	54.3952	69.8577	1.284263287	NO
SKI	41.7994	53.6915	1.284504549	NO
MOBKL1A	6.47008	8.31139	1.284588245	NO
CD2AP	43.5104	55.9909	1.286841172	NO
C4orf41	10.4952	13.5103	1.287290803	NO
NUP133	54.7488	70.5008	1.287713814	NO
C17orf79	57.8105	74.466	1.288104822	NO
LAMA1	9.26114	11.9293	1.288106607	NO
UBE2A	60.1192	77.4734	1.288662078	NO
COG7	12.1892	15.7127	1.289065883	NO
NPDC1	45.2948	58.3947	1.289214214	NO
DISC1,TSNAX	40.0657	51.6646	1.289496628	NO
TMEM165	21.9447	28.3172	1.290393432	NO
LDLR	11.6512	15.0371	1.290604535	NO
TBC1D4	8.25287	10.6514	1.290626005	NO
SLC44A2	25.7778	33.2761	1.290880095	NO
RABGEF1	9.71987	12.5534	1.291517328	NO
EPHA7	39.5087	51.0413	1.291900535	NO
LMBRD1	17.6813	22.8647	1.293159294	NO
NPNT	19.4574	25.1737	1.293785994	NO
CDKN1A	143.127	185.468	1.295830482	NO
KRAS	18.8101	24.3755	1.295876291	NO

HBXIP	60.2853	78.1264	1.295945457	NO
DAB2IP	18.6825	24.2133	1.296044272	NO
ATP8A1	6.88559	8.92763	1.296565419	NO
PCDH9	3.90294	5.06378	1.29742577	NO
HCG11	6.4371	8.35346	1.297704585	NO
FAT4	4.93336	6.40249	1.297795438	NO
TMEM170B	9.84687	12.7793	1.297800836	NO
SCIN	24.3003	31.5531	1.298463986	NO
CEP120	15.0188	19.5164	1.299458896	NO
HSPA12A	6.3296	8.22656	1.299696707	NO
PIK3C3	24.3924	31.7065	1.299849866	NO
SNX25	12.7098	16.522	1.299942671	NO
EFNB1	14.0171	18.2245	1.300162546	NO
GPR173	9.69791	12.6097	1.300249966	NO
NIN	18.3856	23.907	1.300313056	NO
KLF3	13.6155	17.7105	1.300766493	NO
ANKRD12	6.58963	8.57519	1.301315697	NO
COMMD6	13.9736	18.1871	1.301532196	NO
SYPL1	31.4036	40.888	1.302017645	NO
USP6NL	8.4585	11.0155	1.302293836	NO
TMEM50B	16.6795	21.7291	1.302747061	NO
HEG1	6.82928	8.8993	1.303109212	NO
GSK3A	103.674	135.106	1.303181474	NO
SIN3A	24.9277	32.4963	1.303623261	NO
CELF2	13.1521	17.1519	1.304116722	NO
FAM199X	24.4542	31.9073	1.304779481	NO
SMCHD1	34.1257	44.555	1.305615419	NO
VKORC1L1	22.3795	29.2246	1.305860692	NO
RNF220	27.3705	35.7533	1.306269886	NO
ARRDC3	8.47253	11.0707	1.306652942	NO
BCAM	24.3254	31.7863	1.306715437	NO
GOLT1B	15.2706	19.9586	1.306998966	NO
CDH1	6.13807	8.02331	1.307139395	NO
NR2F1	10.9187	14.2791	1.307763805	NO
SEC24D	8.75031	11.4443	1.307869867	NO
NF1	23.6896	31.0075	1.308911904	NO
DDX21	98.8712	129.45	1.309276677	NO
KIF20B	21.7191	28.4509	1.309944783	NO
TUBGCP6	30.6409	40.1396	1.310001079	NO
ADAM9	26.0955	34.1936	1.310324375	NO
SQLE	41.0783	53.8573	1.311087525	NO
DSC2	23.4883	30.7983	1.311218396	NO
CDH11	18.2631	23.9536	1.311588357	NO
PTPN6	8.96642	11.7608	1.311651088	NO
TLE1	27.3893	35.9424	1.312277654	NO
ADAMTS16	5.01947	6.58862	1.31261334	NO
WIPI1	13.4715	17.6915	1.31325675	NO
PITPNM1	17.5976	23.1185	1.313729271	NO
RINL	8.6837	11.4086	1.313793015	NO
OSBPL8	20.4636	26.8894	1.314009768	NO
ZNF711	159.251	209.274	1.314111782	NO
MAN2A1	12.465	16.381	1.314160059	NO
MAGED2	81.1295	106.704	1.315228986	NO
CACNA1G	3.54022	4.65653	1.315319242	NO
C11orf9	3.88031	5.10501	1.315618317	NO
DNAJC6	7.02419	9.24666	1.3164028	NO
MDK	167.606	220.729	1.316952216	NO
CTTNBP2	3.46533	4.56685	1.317867201	NO
LRDD	21.7975	28.742	1.318589045	NO

C9orf72	9.51776	12.5502	1.31860641	NO
ADAM10	19.4132	25.6104	1.319224411	NO
YTHDC2	22.9234	30.2451	1.319398162	NO
TCTN1	14.7335	19.4397	1.319423769	NO
WNT5A	4.37244	5.77042	1.319724692	NO
CXCR4	48.0093	63.4029	1.320637942	NO
FAM60A	26.99	35.6495	1.320839344	NO
WDR90	14.7768	19.5189	1.320913505	NO
LIG4	9.24931	12.2205	1.321231252	NO
RAD21	58.0465	76.6969	1.321301771	NO
MTMR12	14.1419	18.6866	1.321363135	NO
TEX15	2.60382	3.44172	1.321796426	NO
RBL1	11.2492	14.8725	1.32209789	NO
ZNF805	3.20916	4.24465	1.322666185	NO
TICAM2,TMED7,TMED7-TICAM2	32.6526	43.1939	1.322830303	NO
PSD3	7.52646	9.96086	1.32344478	NO
PPA1	52.3896	69.34	1.323545691	NO
RBM6	37.794	50.0239	1.323591562	NO
HEY1	37.8686	50.1236	1.323620004	NO
RNF141	7.0981	9.39741	1.323931978	NO
SPAST	9.37954	12.4232	1.324502898	NO
IMPA1	6.66206	8.82452	1.32459379	NO
C11orf10	91.9277	121.853	1.325526029	NO
NRP2	8.30216	11.005	1.325562781	NO
RASGRP1	2.23108	2.9577	1.325679475	NO
AGPAT5	13.3523	17.7073	1.326167496	NO
SRPX	7.4253	9.84945	1.326470876	NO
ACAP3	20.6755	27.4292	1.326650179	NO
FBN1	4.65553	6.17765	1.32694999	NO
IRS4	193.306	256.705	1.327975012	NO
CCDC99	28.2853	37.5637	1.328030242	NO
RBMS1	12.1833	16.1811	1.328133344	NO
ITGB8	3.74015	4.96797	1.328281568	NO
CACNA2D2	3.1723	4.21445	1.328515445	NO
SPIRE2	11.0629	14.6983	1.328608454	NO
PLXNB2	85.4012	113.55	1.329604343	NO
ZMYND19	61.0817	81.3038	1.3310659	NO
CASP3	23.3569	31.1085	1.331878981	NO
NRP1	15.3001	20.3788	1.331934373	NO
RAD51AP1	26.9873	35.9682	1.332784013	NO
CXADR	18.0587	24.0755	1.333181312	NO
EN2	8.49571	11.3308	1.333707224	NO
MKNK2	53.0321	70.7772	1.334609798	NO
TGFBR2	10.1258	13.5147	1.334683806	NO
TTK	42.0204	56.1266	1.335699062	NO
XPNPEP3	5.17184	6.91005	1.3360926	NO
SECISBP2L	11.217	14.988	1.33618151	NO
PTPRJ	4.18966	5.59893	1.336366757	NO
LEMD3	15.2495	20.3928	1.337269276	NO
DPP8	18.1449	24.2658	1.337329528	NO
NRIP1	8.89038	11.8896	1.33735641	NO
MARCH6	48.5124	64.879	1.337369388	NO
PPP2R2C	6.27701	8.40499	1.339011174	NO
DYNC2LI1	18.9592	25.3919	1.339289643	NO
KLHL14	4.23456	5.67648	1.340512805	NO
DCBLD2	16.6391	22.3185	1.341323288	NO
FMNL2	20.4251	27.4081	1.341887756	NO
CHMP4C	7.28309	9.77585	1.342267301	NO
ZBTB41	3.85969	5.18095	1.342322195	NO

NUP160	22.8559	30.696	1.343021127	NO
CSPG4	1.27514	1.71376	1.343982171	NO
STK38L	10.3675	13.9363	1.344229994	NO
JAG1	11.6717	15.6969	1.344875851	NO
PDE6D	21.1263	28.4254	1.345495905	NO
SCAMP5	6.23516	8.39307	1.346088253	NO
ADM	18.9635	25.5313	1.346342063	NO
TMF1	9.45427	12.7365	1.347171012	NO
IFI30	22.2484	29.9727	1.347181283	NO
PREX1	2.55815	3.44712	1.347504415	NO
KAL1	15.65	21.0942	1.3478706	NO
CHPF	24.3632	32.8425	1.348039714	NO
H3F3B	188.016	253.512	1.34835464	NO
SEC24B	12.8878	17.3842	1.348888407	NO
LRRC4B	3.96969	5.35712	1.349502828	NO
SFRP1	4.62552	6.24246	1.349570179	NO
RIF1	23.2617	31.3943	1.349612275	NO
NAV2	5.75021	7.76103	1.349693664	NO
FZD6	16.3959	22.133	1.349914468	NO
TCEA1	10.8823	14.7095	1.35169626	NO
PPP1R9B	49.4452	66.8527	1.35205515	NO
NCRNA00219	14.0162	18.9529	1.352217291	NO
ARHGAP28	3.94033	5.33088	1.352904496	NO
ERBB4	2.44031	3.3016	1.352942007	NO
SNX16	2.68998	3.64008	1.353200861	NO
IGFN1	0.924076	1.25047	1.353211179	NO
ID4	91.1529	123.385	1.35360237	NO
GRID1	3.48937	4.72811	1.355003896	NO
FAM72B	5.0289	6.81666	1.355498954	NO
GMCL1	14.1826	19.2576	1.357828244	NO
WDR17	1.6998	2.30811	1.357867774	NO
ARHGEF11	14.6406	19.8805	1.357895069	NO
SYTL1	5.55735	7.54874	1.358332807	NO
NPAT	7.74922	10.528	1.358593635	NO
PTK2B	2.2082	3.00137	1.359191749	NO
H3F3C	91.8253	124.814	1.359260526	NO
ODZ4	1.01995	1.3869	1.359782587	NO
ATP6AP2	66.4596	90.3721	1.35980615	NO
TUBGCP5	8.06419	10.967	1.359960737	NO
GFPT2	4.32238	5.88236	1.360906546	NO
PDCD10	24.9297	33.9345	1.361210325	NO
MATN3	3.13754	4.27184	1.36152361	NO
GTPBP2	53.6365	73.0366	1.361695381	NO
NDFIP1	27.2912	37.1626	1.361705763	NO
EFNA3	9.59188	13.0615	1.361721809	NO
GATM	3.47795	4.73609	1.361745406	NO
TTYH2	2.6531	3.61353	1.362005	NO
CDK1	78.9864	107.655	1.3629579	NO
ST14	4.27956	5.83478	1.363405776	NO
NRCAM	11.8781	16.199	1.363770611	NO
DSEL	2.34215	3.19446	1.363899177	NO
CHST6	1.20928	1.65213	1.36621165	NO
RAB14	34.2544	46.8107	1.366563027	NO
PTPRD	10.0269	13.7032	1.366642597	NO
RHPN1	8.21137	11.2225	1.366705119	NO
SSPO	1.5843	2.16584	1.367068941	NO
FBN2	35.5906	48.6732	1.36758547	NO
SLC38A2	72.6865	99.4142	1.367711552	NO
MYST1	21.7679	29.8393	1.370790406	NO

PCDH1	2.72727	3.73857	1.370808459	NO
SBNO2	14.5546	19.9745	1.372389507	NO
LRRC40	13.8647	19.0322	1.372712976	NO
PLEKHA5	15.5378	21.3373	1.373250674	NO
DENND3	3.96528	5.4455	1.373295412	NO
ITGA4	6.55272	9.00028	1.373517222	NO
SNX10	13.268	18.2332	1.374227635	NO
BCL2L11	16.7993	23.0977	1.374920307	NO
ARFGEF2	10.3269	14.1994	1.37498988	NO
RAPGEF4	5.00244	6.88154	1.37563812	NO
EGFL7	18.0345	24.8092	1.375650516	NO
RHPN2	22.866	31.4578	1.375742058	NO
ABHD10	10.3771	14.2775	1.37587366	NO
APCDD1	8.42878	11.6029	1.376578612	NO
TNKS2	31.3937	43.2227	1.376798089	NO
TOPBP1	29.3126	40.3894	1.377884538	NO
CBFB	34.4068	47.4345	1.378637344	NO
DCUN1D4	13.1775	18.1719	1.379011989	NO
COLEC12	6.87543	9.48237	1.379165891	NO
FAM178A	17.3121	23.904	1.380769977	NO
CLK2P	16.8889	23.3203	1.380806346	NO
LTBP1	22.2947	30.7882	1.380964277	NO
ATAD1	30.1472	41.6412	1.381263916	NO
TRAPPC10	22.9535	31.7444	1.382987388	NO
HK2	6.36254	8.79935	1.382992181	NO
PTHLH	6.56992	9.09155	1.383814919	NO
HSP90B3P	21.8147	30.2085	1.384774439	NO
CDYL2	6.67543	9.24616	1.385103708	NO
PIKFYVE	13.0993	18.1571	1.386110238	NO
SLC40A1	7.67617	10.6493	1.387322309	NO
BRMS1L	14.3758	19.9595	1.388403589	NO
ZDHHC21	6.75293	9.37756	1.388665378	NO
TTC32	9.05688	12.5809	1.389095704	NO
CBX3	104.695	145.555	1.390279549	NO
CNOT8	26.37	36.6642	1.390375919	NO
RASA4	1.60711	2.23494	1.390656394	NO
MED17	12.102	16.8366	1.391219443	NO
MALT1	15.3243	21.3237	1.391498159	NO
EHBP1L1	6.16498	8.58153	1.391979534	NO
PCSK5	12.4203	17.3119	1.393841968	NO
SEZ6L2	9.46244	13.1984	1.394825844	NO
KCNK1	20.2933	28.3228	1.395670133	NO
ITGA6	27.4236	38.2743	1.395672067	NO
MAN1A1	22.1148	30.8818	1.39643072	NO
ZHX1	13.101	18.2961	1.396545909	NO
PDHX	15.9744	22.3112	1.396682405	NO
GPR126	20.2629	28.3058	1.396924452	NO
SLC9A2	0.902387	1.26148	1.397933759	NO
C6orf1	13.1204	18.3425	1.398017093	NO
JAG2	8.93012	12.4917	1.398831316	NO
GALNT7	24.9685	34.9469	1.399640193	NO
C19orf51	2.82696	3.95945	1.400601949	NO
FGFR2	11.5061	16.1213	1.401111724	NO
GK5	4.89115	6.85358	1.401219529	NO
ATP2B1	25.8896	36.2804	1.401350654	NO
C8orf73	4.6281	6.48809	1.401888881	NO
KCNH3	2.78891	3.91199	1.402695636	NO
KIAA1143	7.66287	10.7555	1.403580685	NO
ANKRD19	11.386	15.9884	1.404222938	NO

ACVR1C	0.857871	1.20619	1.406028658	NO
SEMA5A	1.52074	2.139	1.406554057	NO
AP1AR	9.28666	13.0693	1.407322527	NO
GREB1	2.52794	3.55893	1.407834748	NO
SEPP1	28.6914	40.4182	1.408722063	NO
RRM2B	13.0071	18.3239	1.40876991	NO
TMEM59L	14.7726	20.8244	1.409663677	NO
CASP8AP2	22.8091	32.159	1.409920679	NO
KIF21B	1.76734	2.49553	1.412021445	NO
FAM72A	6.37664	9.00661	1.412439429	NO
ITGA2	5.57575	7.87713	1.412748835	NO
NCRNA00087	2.08988	2.95429	1.413615731	NO
BOD1L	14.6693	20.7468	1.414302769	NO
UBA2	150.109	212.358	1.414691029	NO
MZT1	10.8783	15.4047	1.416099856	NO
TMEM55B	18.9977	26.9036	1.416150898	NO
C6orf168	4.77466	6.7621	1.416247098	NO
CPNE7	3.63414	5.14771	1.416486645	NO
CACNA2D1	10.361	14.6773	1.416593669	NO
ADAMTS7	5.01583	7.10882	1.417278224	NO
BRP44L	33.5845	47.6041	1.417440327	NO
LBR	85.2356	120.818	1.417461942	NO
KIF27	2.56683	3.63975	1.417990631	NO
PACSIN1	1.87052	2.65279	1.418209829	NO
CD58	6.94691	9.8539	1.418457574	NO
ACSL3	45.7985	65.0325	1.419970541	NO
UBL3	11.0093	15.6368	1.4203259	NO
FREM2	4.55404	6.46905	1.420509027	NO
ANKRD33B	9.31269	13.2308	1.420732554	NO
RASSF6	0.832112	1.18228	1.420817248	NO
ATP1B2	4.52825	6.43386	1.420827096	NO
GCA	4.30164	6.12367	1.423564637	NO
SCN9A	4.81366	6.85619	1.424318706	NO
MYBL1	4.21716	6.00899	1.424890445	NO
NBEAL1	3.40288	4.85236	1.425958505	NO
LRRK2	1.95226	2.78455	1.426323271	NO
ITPR1	2.94883	4.20672	1.426573421	NO
VCAN	50.718	72.4103	1.427704097	NO
SVIP	12.1951	17.429	1.429183338	NO
PCGF2	32.2035	46.0554	1.430135655	NO
NELL2	1.41668	2.02652	1.43047571	NO
ACBD5	15.9549	22.8486	1.432073957	NO
LTBP4	20.6617	29.5951	1.432361851	NO
CD24	33.6158	48.2407	1.435060938	NO
LRP6	12.0199	17.252	1.435285759	NO
LONRF1	8.00208	11.4857	1.435339483	NO
ARPP19	43.4298	62.4493	1.43793753	NO
C2orf89	6.2425	8.97808	1.438218627	NO
YEATS4	25.8383	37.1817	1.439011379	NO
JMJD1C	14.2089	20.4499	1.439235822	NO
TLL1	0.50132	0.721941	1.440080041	NO
GRK4	1.34137	1.93229	1.440531292	NO
SCAPER	7.76518	11.1864	1.440591203	NO
PANK3	27.2677	39.2932	1.441018643	NO
ARHGAP23	6.13036	8.83788	1.44165904	NO
RNF128	11.0765	15.9691	1.441715001	NO
MRPS36	6.25508	9.02264	1.442448689	NO
USP32	14.3779	20.7405	1.442528678	NO
SORBS1	3.60251	5.20234	1.444089341	NO

FAM49A	1.47766	2.13448	1.444506805	NO
ENHO	5.88743	8.50576	1.444732105	NO
BHLHE22	1.67292	2.41963	1.446352296	NO
HIST3H2A	35.8468	51.9067	1.448013447	NO
CILP2	2.21141	3.20511	1.449353988	NO
MSX1	8.30483	12.0492	1.450863707	NO
GRAMD1B	2.70408	3.92441	1.451289164	NO
FAM76B	9.02945	13.1054	1.451408878	NO
SOX11	16.9721	24.636	1.451554761	NO
BTG1	14.114	20.4958	1.452159578	NO
TSPAN4	9.80694	14.2509	1.453149364	NO
ENPP2	9.48993	13.7999	1.454163007	NO
TMEM47	7.1845	10.4481	1.454262798	NO
MGAT4A	5.89822	8.58166	1.454956479	NO
IFNGR2	37.984	55.2714	1.455120874	NO
HAPLN4	0.87994	1.28103	1.455817993	NO
SLC43A2	6.51848	9.4959	1.456766852	NO
ZFAND5	22.5035	32.7876	1.456999114	NO
FAM169A	10.513	15.3219	1.45742536	NO
PSMB10	6.28184	9.1641	1.458824162	NO
ACVRL1	1.02275	1.49278	1.459580721	NO
SSTR2	1.21579	1.77508	1.460018854	NO
LOC100127983	13.0174	19.006	1.460045167	NO
EIF4ENIF1	18.4004	26.8884	1.461295554	NO
BRD7	15.7017	22.9594	1.462223659	NO
SASS6	14.2788	20.897	1.463499243	NO
SORT1	33.8033	49.5034	1.464454124	NO
NPAS1	2.85988	4.18835	1.464520106	NO
LEO1	17.4194	25.5161	1.464811477	NO
UNC5D	1.02031	1.49461	1.464859198	NO
ABAT	4.22324	6.18697	1.464981047	NO
ACAP2	6.87028	10.0665	1.465218681	NO
COG3	5.86965	8.60364	1.465785502	NO
CRIP2	3.92457	5.75558	1.466549737	NO
CCDC61	3.38086	4.95974	1.467005215	NO
BBC3	23.5751	34.5904	1.467242161	NO
C14orf147	14.9396	21.9234	1.467464904	NO
FBN3	0.374728	0.550168	1.468181166	NO
CEP135	7.03067	10.3238	1.468393873	NO
C4orf31	2.20377	3.23714	1.468914068	NO
TMEM128	5.13791	7.54979	1.469427317	NO
POLR2F	29.1874	42.9382	1.471124147	NO
EGF	0.618702	0.910931	1.472325851	NO
MATK	1.62878	2.39874	1.472723915	NO
BAI1	1.90834	2.81433	1.474755715	NO
DAAM1	7.10725	10.4851	1.475271005	NO
CKS2	151.482	223.533	1.475641226	NO
TBK1	14.861	21.9329	1.475870359	NO
BTG2	37.2875	55.0441	1.476206963	NO
ABCA7	7.85985	11.6105	1.477186517	NO
RDH10	7.45991	11.0309	1.478688329	NO
PRODH	2.44792	3.62085	1.479154754	NO
TMEM100	1.3676	2.02373	1.479767994	NO
SBF2	6.81476	10.0864	1.480082916	NO
COL21A1	3.87665	5.73804	1.480156784	NO
SNCAIP	0.760572	1.12578	1.48017833	NO
NPTN	66.9929	99.242	1.481381268	NO
AGA	6.00157	8.89223	1.481649291	NO
KLF15	1.49805	2.22056	1.482296443	NO

UGT3A2	3.29064	4.8786	1.482568743	NO
KCNC3	1.51035	2.23937	1.482680759	NO
CRISPLD1	3.09977	4.59676	1.482935655	NO
SERPINE2	27.0189	40.0682	1.482968548	NO
NCKAP5	0.878279	1.30397	1.484690277	NO
TNFRSF9	0.584292	0.868816	1.486953982	NO
SNX9	30.5083	45.3666	1.487025101	NO
SLC6A15	15.5522	23.1326	1.487416829	NO
ARRDC4	18.5182	27.5637	1.488469851	NO
PRDM1	0.585635	0.87193	1.488861959	NO
ATP1B1	47.3613	70.5284	1.489157141	NO
TSC22D1	24.2886	36.1783	1.489516392	NO
NRG1	1.40284	2.09116	1.490654591	NO
PCDH10	1.03858	1.54961	1.492044951	NO
LRP5	21.419	31.9596	1.492119415	NO
WWC3	12.4015	18.5136	1.492854954	NO
PRSS8	1.42317	2.12634	1.494094086	NO
ITGAV	38.5062	57.6469	1.497083822	NO
HS6ST2	22.1354	33.2645	1.502778103	YES
TCIRG1	2.39658	3.60696	1.50504435	YES
SERPINI1	1.96944	2.96621	1.506115071	YES
CR2	2.64624	3.98587	1.506239308	YES
PTPRK	22.9724	34.633	1.507586729	YES
EML1	13.8332	20.9094	1.511535644	YES
SNAP91	2.89565	4.38518	1.514405966	YES
KIF3A	10.7992	16.3621	1.515118885	YES
FAM107B	12.75	19.3264	1.515792213	YES
NRXN2	0.994234	1.50809	1.516835884	YES
NUDT4	5.59617	8.48875	1.516884249	YES
GABRQ	7.57725	11.4987	1.51753522	YES
ANKRD24	0.600573	0.911521	1.517752974	YES
ADRA2A	0.65766	0.99817	1.517759286	YES
IMPACT	16.8583	25.6003	1.518553777	YES
SLC16A9	4.77382	7.26036	1.520870168	YES
HMGXB4	20.4798	31.1573	1.521362553	YES
ABCG4	1.22308	1.86213	1.522493427	YES
ARL6IP1	142.555	217.233	1.523853276	YES
ARFGEF1	11.5342	17.5854	1.524629821	YES
SPAG1	5.4978	8.38299	1.524790462	YES
IGSF9	1.66118	2.53521	1.526152363	YES
FNDC1	1.84338	2.81412	1.526612596	YES
NAA38	4.61992	7.05751	1.527624539	YES
MTM1	4.70812	7.19597	1.528415719	YES
PRDM8	2.03587	3.11299	1.529069519	YES
CACNA1H	6.77031	10.3562	1.529648317	YES
IRX3	9.67146	14.7959	1.529852964	YES
KLHDC7A	0.373717	0.57176	1.529925073	YES
FAM105A	4.78293	7.32057	1.530564666	YES
CD55	19.8676	30.409	1.53058058	YES
ISCU	39.4779	60.4727	1.531810678	YES
GRIK5	1.78938	2.74171	1.532207831	YES
FNDC5	1.48917	2.28403	1.533763455	YES
ECEL1	6.38454	9.79402	1.534020752	YES
TCF24	2.25647	3.46201	1.534260014	YES
DACT3	2.57474	3.9508	1.534442942	YES
ELAVL2	9.91264	15.2117	1.534579088	YES
PDE10A	5.60491	8.60535	1.535321722	YES
GNAI1	17.3004	26.5704	1.535828367	YES
ADSSL1	2.30031	3.53842	1.538238279	YES

KIAA1432	16.637	25.5981	1.538620036	YES
LAT2	0.858762	1.32239	1.539883278	YES
LOC100130691	0.34735	0.535013	1.540268645	YES
SYT1	0.372862	0.575776	1.544207901	YES
NTRK3	1.81756	2.80814	1.545006599	YES
ADAMTS5	3.64393	5.64105	1.548068167	YES
DNAJB4	8.49969	13.1782	1.550430654	YES
HES2	0.499729	0.774833	1.550506958	YES
SLC25A12	7.45057	11.5548	1.550868109	YES
SH2D3C	1.62719	2.52363	1.550916484	YES
APLP1	11.666	18.1145	1.552761228	YES
SLBP	48.2154	74.9074	1.553598809	YES
CHUK	20.7717	32.3152	1.555728167	YES
HIST1H2BJ	7.95635	12.3905	1.55730983	YES
C1D	7.89582	12.3061	1.558553847	YES
CRTAC1	2.94781	4.59526	1.558869329	YES
PCNA	310.656	485.177	1.561781902	YES
PCDH17	2.53548	3.96184	1.562557197	YES
TTC35	18.5123	28.9291	1.562698004	YES
HERC2P2	17.2627	26.993	1.56365908	YES
IGDCC3	1.10769	1.73307	1.564584958	YES
MED4	15.9905	25.0279	1.565169606	YES
FAM46C	2.10464	3.29742	1.56673589	YES
DLAT	22.4696	35.2438	1.568515733	YES
EPHA3	2.9853	4.68281	1.56862337	YES
PSMG2	39.2133	61.5397	1.569356373	YES
ENTPD2	1.03117	1.61852	1.569597882	YES
CDHR1	3.57333	5.61429	1.571162063	YES
GUCY1B3	7.17235	11.2819	1.572966559	YES
SOSTDC1	1.10207	1.73444	1.573793222	YES
UBE2B	16.4005	25.8175	1.574184893	YES
CNTN4	5.2036	8.19554	1.574972895	YES
WDR19	7.57247	11.9281	1.575194523	YES
GBGT1	7.87788	12.4117	1.575517741	YES
GDF6	1.17798	1.85603	1.575601833	YES
RAC2	1.68538	2.65897	1.577668393	YES
TBX1	15.3336	24.2195	1.579508828	YES
SH3TC1	0.566849	0.896079	1.580806737	YES
UNC5A	1.36545	2.15997	1.581872147	YES
C6orf154	1.52847	2.41893	1.582588304	YES
RTN4RL1	1.81022	2.86518	1.582783576	YES
LGR4	19.0485	30.1516	1.582881221	YES
DAGLA	3.76117	5.96685	1.586434547	YES
FOLR1	4.33191	6.87452	1.586949259	YES
SCUBE3	1.99443	3.16594	1.587391516	YES
HSPA13	29.0183	46.0658	1.587478442	YES
ACHE	1.88151	2.98954	1.588902941	YES
RAPGEF5	6.0183	9.56998	1.590146845	YES
ABCG1	1.8924	3.01559	1.593525383	YES
CYP26A1	1.38919	2.2148	1.594309805	YES
MAP9	7.21855	11.5144	1.595116726	YES
MAK	1.24533	1.98684	1.595427444	YES
IARS2	66.872	106.705	1.595653057	YES
GATSL3	1.99917	3.19035	1.595835561	YES
CCKBR	4.3499	6.95458	1.598790605	YES
FGD4	3.48682	5.57863	1.599919148	YES
LOC100329109	6.33081	10.1375	1.601294873	YES
MICA	1.58342	2.53672	1.602056468	YES
CCNA1	4.27296	6.84946	1.602978416	YES

CLCNKB	0.808868	1.29688	1.603328451	YES
GCLC	23.9971	38.4966	1.604219995	YES
ITGA11	0.32663	0.524028	1.604345652	YES
TLR3	0.637834	1.02393	1.605326778	YES
FOXA3	0.829804	1.33261	1.605938895	YES
TERT	1.39822	2.24723	1.607215074	YES
LRRN1	0.951599	1.53096	1.608830124	YES
PSMG4	21.9432	35.3224	1.609715801	YES
PRRX1	0.43204	0.697101	1.613509419	YES
FST	4.41621	7.12699	1.61382372	YES
KNDC1	0.354271	0.571806	1.61403739	YES
SLC4A7	4.70766	7.60055	1.61450734	YES
PRKCD	14.2158	22.952	1.614542032	YES
TOMM6	149.036	240.639	1.614638279	YES
ARHGEF4	1.81691	2.94601	1.621442653	YES
ISCA2	15.3577	24.907	1.621787727	YES
HIAT1	25.5897	41.5375	1.623212635	YES
RAB3GAP2	12.5756	20.452	1.626328844	YES
C13orf36	0.431412	0.70242	1.628186543	YES
DEK	175.007	285.047	1.628779151	YES
LIN54	6.05493	9.86267	1.628867215	YES
GPCPD1	6.05462	9.86948	1.630073481	YES
PDE1B	2.18835	3.57011	1.63141407	YES
NCKAP1L	0.877022	1.43222	1.63304325	YES
SHISA2	4.46196	7.29149	1.634146131	YES
IGF2,INS,INS-IGF2	9.0784	14.8432	1.635004946	YES
PLIN1	0.516527	0.846065	1.637987106	YES
TRAPPC4	28.0584	45.9644	1.638168774	YES
SEMA3F	1.95823	3.21199	1.640251454	YES
LPL	4.22397	6.9291	1.640424277	YES
RABGGTB	29.6157	48.6787	1.643679481	YES
EMR2	0.523552	0.860622	1.643813926	YES
PGM2L1	6.08158	10.0132	1.646485721	YES
PCSK9	0.754191	1.24278	1.647824959	YES
PPP1R1A	2.65719	4.38006	1.64837787	YES
DNER	3.44393	5.71166	1.658468948	YES
SOX8	1.73592	2.88268	1.660605061	YES
MOSPD2	10.2667	17.0702	1.66267708	YES
ADAMTS18	0.668308	1.11119	1.662689757	YES
DCK	12.3201	20.5084	1.664637451	YES
PAIP1	9.72504	16.194	1.665186768	YES
RHBDL3	3.94337	6.57569	1.667529172	YES
SLC7A11	1.47672	2.46685	1.670494231	YES
ICAM5	4.19121	7.00794	1.672056966	YES
TPPP3	2.83794	4.75908	1.676951541	YES
RECK	6.55997	11.0076	1.678001494	YES
TBC1D14	17.8693	30.0043	1.679100985	YES
RGS5	1.11266	1.8685	1.679319806	YES
CCNG2	8.45252	14.2329	1.683863275	YES
SHANK1	0.794915	1.34066	1.686551064	YES
RGS17	0.990327	1.67541	1.691776515	YES
MEGF9	11.2773	19.0896	1.692745401	YES
MMP11	1.74691	2.95899	1.693845161	YES
NOP16	22.9517	38.9454	1.696841723	YES
NGFR	3.62375	6.16871	1.702298456	YES
DES	0.998975	1.70767	1.70942248	YES
STXBP3	17.1061	29.2762	1.71144627	YES
PNRC2	2.24438	3.84316	1.712352831	YES
ATP13A3	22.2267	38.0922	1.713801476	YES

FGFR3	6.00752	10.2986	1.714277897	YES
ATRN	23.1413	39.7276	1.716744094	YES
LYPD6	5.80253	9.96219	1.716869043	YES
ARL8A	23.8238	41.036	1.722479723	YES
CITED1	5.12072	8.83745	1.725821194	YES
EDA	3.68034	6.3635	1.72905289	YES
GYLTL1B	4.25476	7.367	1.731473131	YES
ZNF706	6.96202	12.0557	1.73163156	YES
CDR1	1.25856	2.17952	1.731751592	YES
LAMC3	2.12046	3.68968	1.740040638	YES
PTPRB	1.20376	2.09508	1.740448349	YES
HIST2H2BA	2.75846	4.81175	1.744359017	YES
CCDC64	2.55873	4.47735	1.749830282	YES
C7orf57	0.951494	1.66553	1.750434405	YES
DOCK9	5.9287	10.4055	1.755103366	YES
KIF21A	11.38	20.0544	1.76225659	YES
CALM2	201.369	354.957	1.762718378	YES
CBFA2T3	0.338624	0.597459	1.764373503	YES
CRLF1	7.54655	13.3374	1.767351501	YES
NOV	4.71945	8.34732	1.768705683	YES
NPY	3.3024	5.85779	1.773797109	YES
MND1	12.1457	21.5563	1.774810508	YES
MEGF6	0.951772	1.69109	1.776776233	YES
VAV3	2.98402	5.30396	1.777456188	YES
ITGAX	0.62203	1.10677	1.779287941	YES
AK7	0.517809	0.921616	1.779838082	YES
CA2	11.8034	21.0316	1.781827901	YES
GNL3	59.0684	105.315	1.782938574	YES
NECAB1	1.28405	2.29307	1.785805551	YES
ATP1A3	3.80968	6.8218	1.790647051	YES
NTS	1.17519	2.10783	1.793613448	YES
C6orf165	0.668533	1.20195	1.797884074	YES
SLIT2	31.5087	56.7033	1.79960964	YES
VSTM2B	3.21864	5.80245	1.802767067	YES
ELFN1	0.661049	1.19228	1.803619484	YES
HIST1H1C	39.7552	71.7332	1.804370995	YES
L1CAM	2.84601	5.15661	1.811873206	YES
DLX5	2.11959	3.84109	1.812187207	YES
LRP2	2.42048	4.40434	1.819617229	YES
ITGB6	0.369844	0.681375	1.842330637	YES
HHEX	9.00722	16.6029	1.843286085	YES
MMP17	1.87061	3.45291	1.845877739	YES
ADAMTS9	1.80938	3.34627	1.849406033	YES
SEMA3D	3.05836	5.65676	1.849607304	YES
EBF2	0.628075	1.16664	1.857481703	YES
SFRP2	1.72042	3.1993	1.859595706	YES
C8orf4	0.678124	1.26197	1.86096897	YES
EBF1	0.649926	1.21414	1.868114651	YES
PVRL4	5.82453	10.9002	1.871428591	YES
CHRM4	0.769352	1.44378	1.876620591	YES
DCLK3	0.534729	1.00981	1.888449232	YES
PRKCG	0.418687	0.791189	1.889690547	YES
TSPAN2	1.62995	3.08321	1.891603872	YES
SHISA3	0.756324	1.43495	1.897272626	YES
NMUR2	1.11723	2.12017	1.897697448	YES
LGR5	12.8508	24.3896	1.897903974	YES
SARDH	0.50834	0.965381	1.899087001	YES
CDK7	12.5282	23.859	1.904421736	YES
ANPEP	0.382604	0.728977	1.90530505	YES

CTNND2	0.417249	0.79516	1.905721103	YES
ISLR2	0.517278	0.988696	1.911343397	YES
CFI	1.25676	2.40311	1.912155698	YES
GABRR2	0.478872	0.922323	1.926032437	YES
CDKN1C	2.16335	4.18335	1.933736205	YES
THBS2	1.42656	2.76014	1.934826228	YES
RASGRP2	0.686864	1.33149	1.938512449	YES
PDZRN3	1.15813	2.2494	1.94226762	YES
EDEM3	10.9235	21.3211	1.951849749	YES
LRAT	0.266461	0.521448	1.956942003	YES
TAS2R5	0.802419	1.57297	1.960287149	YES
GRHL1	5.92872	11.6604	1.966757384	YES
UNC50	24.6037	48.4393	1.968785585	YES
ONECUT3	1.05538	2.07957	1.970445709	YES
SEL1L	18.3655	36.377	1.980726931	YES
TJP3	0.528191	1.04972	1.98739274	YES
PCDH20	0.729436	1.455	1.994690621	YES
WNT11	4.4967	8.99786	2.000984511	YES
STARD8	0.93483	1.87808	2.009017318	YES
FAM78B	0.404337	0.814028	2.013241162	YES
NOVA2	1.5606	3.14753	2.016872662	YES
NOTCH3	13.1609	26.6222	2.022822877	YES
SNCG	1.20291	2.45634	2.041995943	YES
MAP4K3	11.6166	23.7843	2.047438333	YES
ARHGEF10L	5.18394	10.7628	2.076177802	YES
PLAT	16.1172	33.5528	2.081797872	YES
CPXM1	8.69207	18.2024	2.094143057	YES
RET	0.800883	1.69499	2.116396592	YES
ADAMTS3	2.53918	5.39984	2.126601921	YES
WNT3A	0.781171	1.66157	2.127014694	YES
PNRC1	33.2197	70.7549	2.129906353	YES
PCDH19	0.993341	2.12932	2.143591499	YES
PLIN5	0.265511	0.572607	2.15661702	YES
CHP2	0.678583	1.46754	2.16266469	YES
CLVS2	0.335429	0.73028	2.177148779	YES
ALDH1A3	9.82035	21.5886	2.198348329	YES
PHACTR1	0.406072	0.893475	2.20028438	YES
ERLEC1	23.9708	52.969	2.209729844	YES
ADCYAP1	0.341201	0.758738	2.223727487	YES
CCDC12	18.7211	41.9351	2.240001968	YES
VLDLR	22.1755	49.996	2.2545664	YES
CREB3L1	1.25	2.87154	2.297237472	YES
CCDC85A	0.326896	0.752217	2.301094129	YES
KDR	0.409105	0.955675	2.336015748	YES
DGKK	0.405859	0.952339	2.346482987	YES
FZD8	7.80374	18.5671	2.37925516	YES
SCUBE1	0.391749	0.933083	2.381829263	YES
ANO4	0.60233	1.45329	2.412786063	YES
SMPDL3A	5.59529	13.9115	2.486299338	YES
GOLIM4	54.9989	138.864	2.524854059	YES
DNASE1L2	4.57875	11.6706	2.548856231	YES
DOCK10	0.258363	0.662671	2.564895414	YES
LEFTY2	0.294947	1.12481	3.813617101	YES

Table S1. ***Complete list of PUM regulated RNAs in RNA –Seq analysis.***

Table S2

Category	Term	Count	%	PValue
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	PSYCH	41	5.899280576	0.005833855
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	DEVELOPMENTAL	19	2.73381295	0.074206641
OMIM_DISEASE	Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans	6	0.863309353	0.010178811
OMIM_DISEASE	Newly identified loci that influence lipid concentrations and risk of coronary artery disease	5	0.71942446	0.02629997
OMIM_DISEASE	Common variants at 30 loci contribute to polygenic dyslipidemia	6	0.863309353	0.031560847

Table S2. ***DAVID GO analysis of PUM repressed targets and disease associations.***