Ribonomic and Mechanistic Analysis of the Human Pum1 RNA Binding Protein

by

Adam Remy Morris

University Program in Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Jack Keene, Chair

_____
David Pickup

_____
Jen-Tsan Chi

_____
Uwe Ohler

_____
Dennis Thiele

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the University Program in Genetics and Genomics in the Graduate School
of Duke University

2010

ABSTRACT

Ribonomic and Mechanistic Analysis of the Human Pum1 RNA Binding Protein

by

Adam Remy Morris

University Program in Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Jack Keene, Chair

_____
David Pickup

_____
Jen-Tsan Chi

_____
Uwe Ohler

_____
Dennis Thiele

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the University Program in Genetics and Genomics in the
Graduate School
of Duke University

2010

Abstract

Much of the regulation of gene expression occurs at the posttranscriptional level, and much of this regulation is controlled and coordinated by RNA binding proteins (RBPs). Many RBPs have multiple mRNA targets, and the proteins encoded by these targets often share functional relationships, forming posttranscriptional RNA operons. These operons often reflect the function of the RBP, thus determination of the genome-wide targets of RBPs allows insight into their functions.

The PUF family of RBPs is characterized by the presence of an extremely well conserved RNA binding domain, typically consisting of 8 repeats of an RNA binding motif, with each repeat binding to one RNA base. PUF proteins are proposed to have an ancestral role in self-renewal of stem cells and have been shown to affect a number of developmental processes. Human and other vertebrate genomes contain two canonical PUF genes, Pum1 and Pum2, and at the outset of this study there was very little known about functions or targets of either protein, especially Pum1.

In order to identify the genome-wide targets of human Pum1 we used RNA immunoprecipitation followed by microarray, or RIP-Chip, analysis. RIP-Chip allowed us to identify Pum1 target mRNAs in human HeLa cells. We found that there were numerous functional relationships among the proteins encoded by these mRNAs, forming putative RNA operons. Some of these potential operons are progression of cell

cycle, cell differentiation and proliferation, and regulation of transcription. We were also able to find a consensus Pum1 binding motif, UGUAHAUA, in the 3' UTRs of Pum1 target mRNAs.

The genome-wide targets of PUF proteins from other species have been previously identified, and by comparing the targets of human Pum1 to targets of Drosophila Pumilio and yeast Puf3, both of which bind to the same RNA sequence as Pum1, we determined that there has been evolutionary rewiring of regulation by Puf proteins. While the PUF RNA binding domain and consensus binding sequence have remained almost identical through evolution, the surrounding protein sequence and the mRNAs bound have changed dramatically, indicating that evolutionary rewiring is occurring in a modular fashion.

After identifying Pum1 associated mRNAs, we went on the study the function of Pum1. Through Pum1 knockdown assays we found that Pum1 enhances decay of target mRNAs, and that this effect is likely due to Pum1 enhancing deadenylation of these mRNAs. We also showed by immunofluorescence that Pum1 protein has a cytoplasmic granular subcellular localization and upon oxidative stress relocates to stress granules but not processing bodies. We were, however, unable to detect any difference in Pum1 mRNA targeting after stress. We were also unable to detect any changes in progression through cell cycle after Pum1 knockdown.

In this study we identified the genome-wide mRNAs associated with Pum1, determined functional relationships among these targets related to the proposed ancestral role of PUF proteins in self-renewal of stem cells, and identified a sequence motif to which Pum1 binds in these mRNAs. We also demonstrated that Pum1 enhances decay of associated mRNAs, and that this effect is likely due to Pum1 enhancing deadenylation of associated mRNAs. These results provide a description of mRNA targets and mechanisms of action of Pum1 proteins, which will provide a strong foundation for future experiments to further explore the functions of the Pum1, especially as they relate to human stem cells.

Contents

List of Tables

List of Figures

List of Abbreviations

β2M- Beta-2-Microglobulin

Act. D- Actinomycin D

B2M- Beta-2-Microglobulin

FACS- Fluorescence-activated Cell Sorting

FBF- Fem Binding Factor

GI- Guanine and Inosine

GMM- Gaussian Mixture Modeling

GO- Gene Ontology

IF- Immunofluorescence

IP- Immunoprecipitation

KD- Knockdown

LOD- Log Of Odds

miRNA- Micro Ribonucleic Acid

NGS- Normal Goat Serum

OC- Oligo Control

OE- Overexpression

PAT- Poly Adenosine Tail

PCNA- Proliferating Cell Nuclear Antigen

PCR- Polymerase Chain Reaction

PTR- Posttranscriptional Regulation

PTRO- Posttranscriptional RNA Operon

PUF- Pumilio and FBF

qPCR- Quantitative Polymerase Chain Reaction

RBP- RNA Binding Protein

RIP- RNA binding protein Immunoprecipitation

RIP-Chip- RNA binding protein Immunoprecipitation followed by microarray

RNP- Ribonucleoprotein

RT- Reverse Transcription

SEM- Standard Error of the Mean

siRNA- Short Interfering Ribonucleic Acid

SLBP- Stem Loop Binding Protein

UTR- Untranslated Region

USER- Untranslated Sequence Element for Regulation

VC- Vector Control

Acknowledgements

I would like to thank Neelanjan Mukherjee for major contributions to Morris, Mukherjee et al. 2008 and Morris, Mukherjee et al. 2009 and major intellectual contribution other work, Kyle Mansfield, Matt Friedersdorf, and Marshall Thompson for their major intellectual contributions to this work, Shelton Bradrick for helpful suggestions regarding Morris, Mukherjee et al. 2008, Nancy Kedersha for the gift of fluorescently labeled stress granule and p-body markers, the laboratories of Haifan Lin, Matthias Gromeier, David Pickup, Ken Alexander, Blanche Capel, Elwood Linney, and Margaret Kirby for sharing of reagents and equipment, all of my committee members for their guidance and suggestions, Jack Keene for being a great advisor and helping me to become an independent scientist, my sister for being very supportive and giving me an excuse to travel around the world to visit her, and most of all I would like to thank my parents, who have always encouraged me to pursue my own path in life and have constantly supported me while I was doing so.

1. Introduction

1.1 Posttranscriptional Regulation of Gene Expression

Eukaryotic gene expression is a complex process that integrates myriad signals to coordinate production of thousands of gene products in effectively precise temporal and spatial pattern (Niehrs and Pollet 1999; Maniatis and Reed 2002; Orphanides and Reinberg 2002; Hieronymus and Silver 2004). It is generally believed that improper regulation of gene expression can lead to many human defects and disorders. Thus, understanding the underlying mechanisms of gene expression has become an active subfield of genetic diseases and medicine. For example, the advent of microarray technology was crucial for the study of global gene expression and has led to advances in understanding and categorizing human disease (Perou, Jeffrey et al. 1999). As discussed below, while traditional microarray approaches have limitations for understanding some aspects of gene expression, such as transcription rates, novel applications have emerged that provide insight into underlying mechanisms of gene coordination at the posttranscriptional level (Tenenbaum, Carson et al. 2000; Wang, Liu et al. 2002; Grigull, Mnaimneh et al. 2004; Hieronymus and Silver 2004).

Many studies of global gene expression focus solely on the transcriptome, and the only factor assessed is mRNA abundance, which is but a single aspect of gene expression. Several studies have suggested that measuring global mRNA levels to assess the transcriptome using microarrays can be misleading, as gene expression has

multiple layers that manifest themselves in both the nucleus and the cytoplasm after transcription has ended (Gygi, Rochon et al. 1999; Ideker, Thorsson et al. 2001; Keene 2001; Maniatis and Reed 2002; Wilusz and Wilusz 2007). A prime example of the importance of the post-transcriptional environment was recently demonstrated in a study from the laboratory of David Baltimore, which showed that mRNA stability strongly influences gene expression induction kinetics during the inflammatory response, in some cases overriding the effects of transcription (Hao and Baltimore 2009). Regulation of transcription is also a relatively slow process for responding to cellular perturbations; in the sea urchin embryo the average gene is translated seven times more rapidly than it is transcribed (56 minutes for transcription versus 8 minutes for translation) (14) (Ben-Tabou de-Leon and Davidson 2009). In addition, some human genes can take as long as 16 hours to be transcribed (15) (Tennyson, Klamut et al. 1995). In order to rapidly produce a protein it is advantageous to increase the rate of translation immediately, without waiting for transcription and subsequent mRNA export and cytoplasmic regulation (11, 16) (Keene 2001; Mansfield and Keene 2009). Even in the case of genes that are shorter and thus transcribed more quickly, translational control of existing mRNA provides a rapid means to generate necessary proteins in response to cellular signals. Moreover, repression of transcription can be a relatively slow means to cease protein production if not coupled to rapid mRNA decay and/or a decrease in translation. In addition, many studies have shown that

transcription may be more stochastic than was previously believed, and therefore the newly synthesized mRNA populations created by transcription may be altered correspondingly in the posttranscriptional environment (Kaplan, Kahn et al. 1992; Blake, M et al. 2003; Rodriguez-Trelles, Tarrio et al. 2005; Raj, Peskin et al. 2006; Yanai, Korbel et al. 2006). The underlying mechanisms of posttranscriptional regulation (PTR) are determined by many factors that may bind to and regulate an mRNA after transcription and up to and during translation, including RNA-binding proteins and small noncoding RNAs, such as microRNAs. Recent studies have shown that understanding PTR on a global level provides insights into the coordination of gene expression and its implications for disease (Keene 2007; Halbeisen, Galgano et al. 2008; Mansfield and Keene 2009).

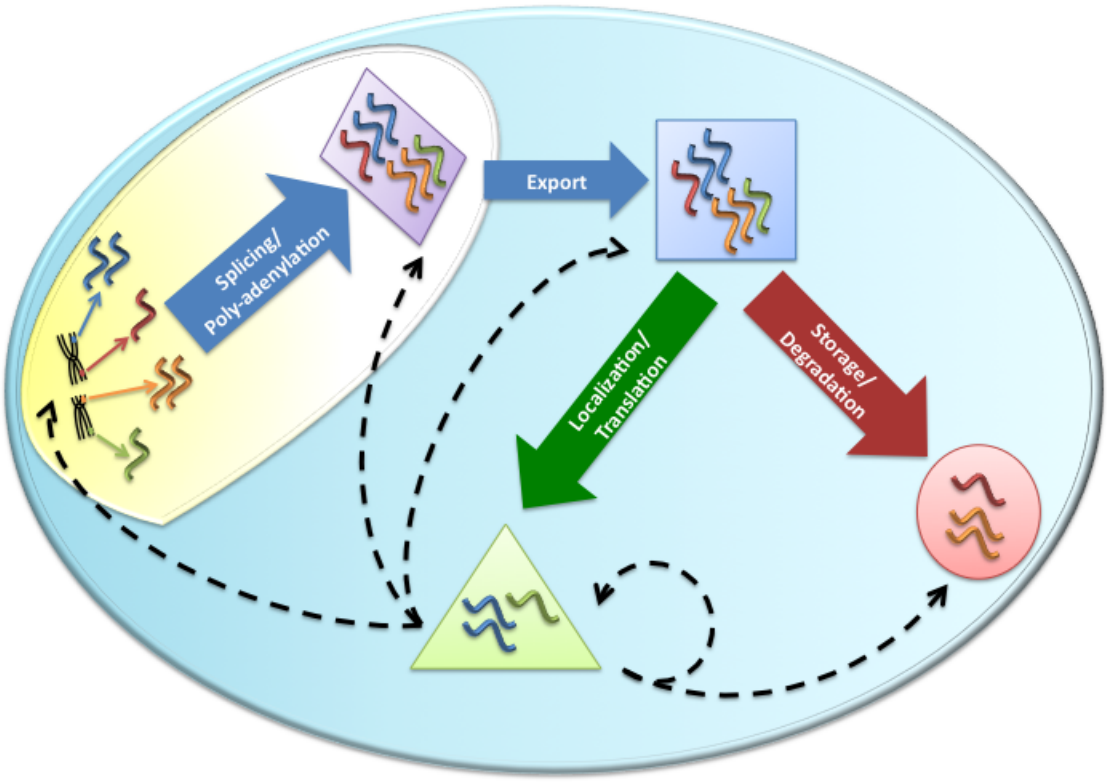


Figure 1: Coordination of post-transcriptional regulation. New transcripts (squiggled lines) emerge from chromosomal DNA and undergo multiple interconnected steps of regulation from splicing through translation. RBPs coordinately regulate functionally related sub-populations of mRNAs existing in the same state as depicted within different colored shapes, each representing a unique combination of trans-acting factors (e.g. RBPs and microRNAs). The dotted lines depict the 'regulators of regulators' concept presented in the text.

The life of every copy of an mRNA involves multiple points for regulation (Figure 1), including splicing, polyadenylation, transport from the nucleus, localization within the cytoplasm, translation, and decay (for a review of the life of an mRNA we recommend (Moore 2005; Keene 2007; Halbeisen, Galgano et al. 2008; Mansfield and Keene 2009)). In most of the cases where PTR has been shown to be important, the

regulation involved one or multiple RNA binding proteins. Thus, RNA binding proteins have key roles in post-transcriptional gene expression, coordinating many aspects of the life of an mRNA (Moore 2005; Keene 2007).

Regulation of gene expression at the post-transcriptional level involves both control and coordination. While control describes an individual interaction that results in a specific outcome, coordination describes a process of integrating multiple control functions to achieve a higher level of harmonized outcome (Mesarovic, Sreenath et al. 2004). Historically, the study of PTR has focused on the one-on-one small-scale control functions, which may result in profound outcomes, but does not address overall RNA coordination. For example, an RBP or microRNA may affect an mRNA sequence element within a Luciferase reporter, demonstrating control of the expression of that RNA. In addition, those same trans-acting factors may also be found to alter the phenotype of a cell or organism, and could be presumed to do so by affecting the same single mRNA target from which the sequence present in the reporter system was derived in vivo. However, the phenotypic change is just as likely to result from the combined effects of that trans-acting RBP or microRNA on coordinating multiple mRNA targets. In addressing this issue, advances in molecular biological techniques and detection methods have allowed study of control on a wider basis, often global, thus leading to a greater understanding of RNA coordination. Gao et al. (1994) found that the ELAV/HuB RBP can target multiple mRNAs in vitro using total brain mRNA and

suggested that this could represent a coordinating function for posttranscriptional gene regulation (Gao, Carson et al. 1994). Subsequently, Tenenbaum et al. (2000) demonstrated that the HuR and HuB RBPs target multiple mRNAs in vivo during neuronal differentiation in mouse embryonic carcinoma P19 cells (Tenenbaum, Carson et al. 2000). Similar multiple targeting interactions by microRNAs were predicted using computational algorithms, and it is generally assumed today that microRNAs, like RBPs, can target and affect multiple mRNAs in living cells (27-29) (Lai 2002; Lewis, Shih et al. 2003; Friedman, Farh et al. 2009). While RBPs have been shown in many studies to target functionally related mRNAs, such conclusions have not emerged for microRNAs. This is consistent with the very broad target predictions of microRNAs; however, microRNAs have been reported to have profound effects on phenotypes. (Lim, Lau et al. 2005). Interestingly, the actions of specific microRNAs have been shown to fine-tune the production of multiple proteins, possibly acting as a multi-targeted mRNA rheostat, but functional coherence has not been demonstrated among the affected proteins (Baek, Villen et al. 2008; Selbach, Schwanhausser et al. 2008). Indeed, coordination of PTR has best been demonstrated by identifying the genome-wide mRNAs associated with particular RBPs in RNP particles.

1.2 Ribonomics- global analysis of RNPs

The global analysis of mRNA and protein components of RNPs has been termed ribonomics because it explores the ribonome, the total RNP content of a cell including

proteins, mRNAs, and noncoding regulatory RNAs (Tenenbaum, Lager et al. 2002). The majority of ribonomic experiments have employed Ribonucleoprotein-ImmunoPrecipitation-microarray (RIP-Chip) and more recently RIP-Seq, when deep sequencing procedures are employed in place of microarrays (Keene, Komisarow et al. 2006). RIP-Chip involves immunoprecipitation of RNP complexes, typically through use of an antibody to one of the RBP components, extraction of the associated RNAs, and identification of this RNA population on a microarray (Figure 2). Many variations of this method have been employed, such as using a recombinant protein with a physical (e.g. epitope) tag as well as an antibody against the endogenous protein for immunoprecipitation (Tenenbaum, Carson et al. 2000; Keene, Komisarow et al. 2006). RIP-Chip has proven successful in various tissues and species, from yeast to mammals (Morris, Mukherjee et al. 2009). One of the main benefits of RIP-Chip is that it is designed to recover entire RNP complexes, allowing identification of components other than mRNAs, such as other regulatory or RNA processing proteins, as well as small noncoding RNAs. A recent study by Tuschl and colleagues encompassed an in-depth characterization of protein and mRNA components of Ago RNPs (Landthaler, Gaidatzis et al. 2008), demonstrating the usefulness of RIP-Chip type experiments for identifying both the protein and RNA components of RNPs. Other recent advances made by our lab in RIP-Chip analysis, as discussed below, also demonstrate that RIP-Chip is an ideal

method for studying remodeling of RNPs during the dynamic processes of posttranscriptional gene expression (Mukherjee, Lager et al. 2009).

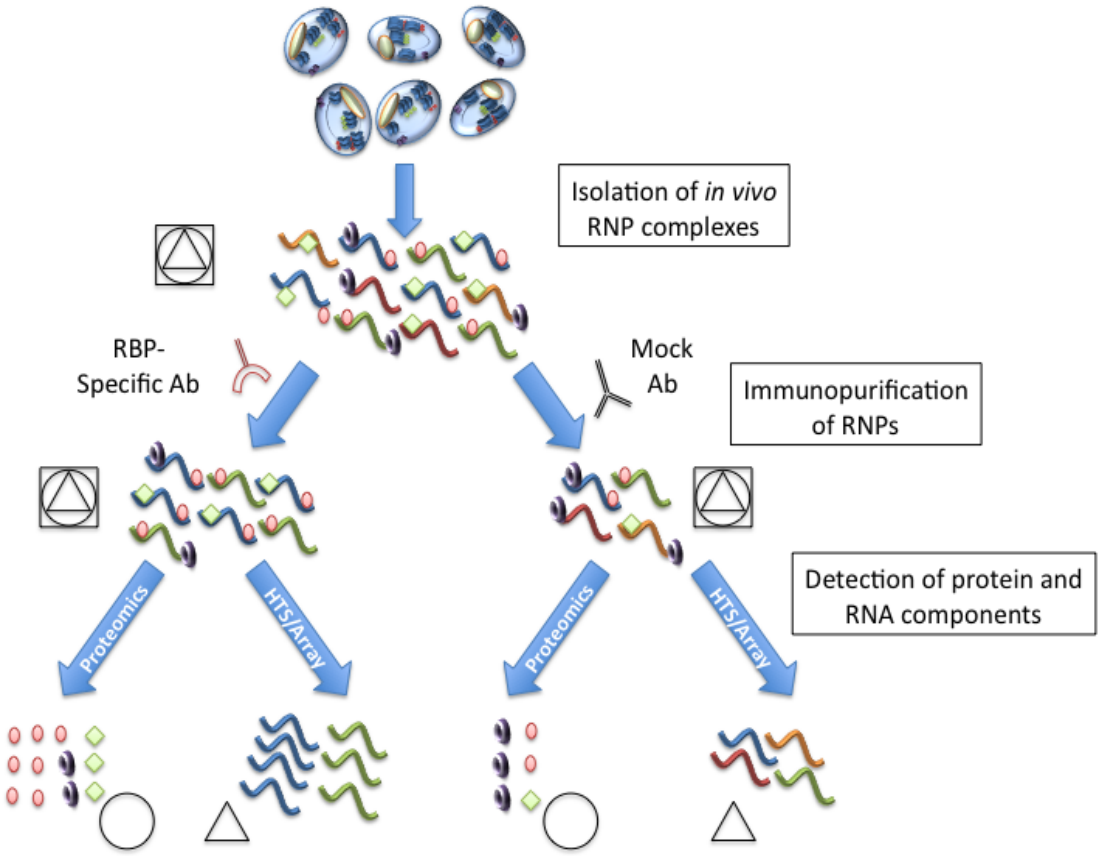


Figure 2: Ribonomics Overview. Tissues or cells are lysed with conditions optimized to preserve in vivo RNP complexes (triangle-RNA inscribed in a circle-protein inscribed in a square-complex). RNPs of interest are immunopurified with an antibody to a specific RBP component. In parallel, a mock IP is performed as a control. Proteins (circle) and mRNAs or microRNAs (triangle) enriched versus the mock immunopurification are detected using proteomics and/or microarray or high-throughput sequencing.

A less commonly used adaptation of RIP-Chip based upon a step of UV-cross linking prior to immunoprecipitation, termed CLIP (cross-linking and immunoprecipitation) or HITS (high-throughput sequencing) -CLIP, can also be used to

determine global mRNAs associated with RNA binding proteins (Ule, Jensen et al. 2003; Licatalosi, Mele et al. 2008). The various modifications of CLIP are designed to identify the specific RNA sequences bound to an RBP of interest, through either cloning or more recently through deep sequencing. UV irradiation forms covalent bonds (albeit inefficiently) between proteins and RNAs that are in direct contact, and thus the RNA that was not in direct contact can be digested away after cross-linking, allowing identification of the exact interacting site. The use of cross-linking also allows protein-RNA interactions to be retained during the extensive purification steps of CLIP. After UV treatment and immunoprecipitation, RNA-protein complexes are separated by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose. Then, the protein component is digested away with proteinase K and the remaining RNA is cloned through use of linker ligation and RT-PCR or high throughput sequencing (Ule, Jensen et al. 2003; Licatalosi, Mele et al. 2008).

Ribonomics analysis has led to many insights into both control and coordination of PTR. As the true value of the ribonomics approach is its global nature, the remainder of this section will focus on the insights that pertain to PTR on a global level.

(i) One key aspect in understanding how RBPs can coordinate PTR is the insight that RBPs typically associate with multiple mRNAs. The range of mRNAs that a typical RBP associates with is very wide, from a large majority of mRNA species to only a few; however it appears that most RBPs are associated with multiple species of mRNA. This

concept was demonstrated by Gao et al. and Tenenbaum et al. (Gao, Carson et al. 1994; Tenenbaum, Carson et al. 2000), and while confirmed in dozens of published studies, was recently supported most convincingly by Patrick Brown's laboratory in a ribonomic study of 46 RBPs in yeast, which showed that almost all of the RBPs studied were associated with multiple species of mRNA (Hogan, Riordan et al. 2008).

(ii) When the multiple mRNAs associated with an RBP are analyzed, it is typically found that the encoded proteins share a functional relationship. This is another property of the posttranscriptional environment that has been revealed through ribonomic analysis; the mRNAs associated with an RBP are functionally related, for example in encoding a macromolecular complex, signaling cascade, or developmental process (Keene 2001). This property, arising from the discovery that RBPs are multi-targeted (Gao, Carson et al. 1994), contributed to the proposal of the RNA operon/regulon model (Keene and Tenenbaum 2002; Keene and Lager 2005; Keene 2007). This model, demonstrated in Figure 1, describes how RBPs coordinate PTR by associating with multiple, functionally related mRNAs and coordinating their protein production or RNA stability, analogous to the way DNA operons allow coordinated protein production in bacteria. The PTRO model proffered the coordinated spatial and temporal production of functionally related proteins, even when the genes encoding these proteins are found dispersed throughout the genome. RNA operons also

accommodate the multi-functionality of eukaryotic proteins, allowing a single gene to participate in multiple operons.

A compelling example of a mammalian PTRO with a likely role in balancing inflammation was discovered by Paul Fox and Barsanjit Mazumder (Mazumder, Sampath et al. 2003; Mukhopadhyay, Ray et al. 2008; Vyas, Chaudhuri et al. 2009). An RNP known as the GAIT (IFN-gamma-activated inhibitor of translation) complex that contains a phosphorylated form of ribosomal protein L13a binds to and reduces translation of the mRNAs encoding the inflammation protein ceruloplasmin (Cp), VEGF, ZIP kinase and DAP kinase (Mukhopadhyay, Ray et al. 2008). It was hypothesized that translation of other mRNAs encoding pro-inflammatory proteins was also silenced by the GAIT complex, a putative posttranscriptional operon (Vyas, Chaudhuri et al. 2009). To test this hypothesis, the authors used genome-wide translational profiling to discover an array of numerous chemokine mRNAs encoding proteins involved in inflammation that are silenced in a similar manner to the Cp mRNA (Vyas, Chaudhuri et al. 2009). As predicted, many of these mRNAs involved in the GAIT PTRO contain a cis hairpin element, which was shown to mediate silencing of translation. These mRNAs were also responsive to knockdown of the ribosomal L13a protein, a key component of the GAIT complex whose phosphorylation by the ZIP kinase is the switch to silence translation of these functionally related mRNAs (Mukhopadhyay, Ray et al. 2008). Interestingly, the ZIP kinase itself is activated by

phosphorylation by the other GAIT targeted protein kinase, DAP (Mukhopadhyay, Ray et al. 2008). The authors postulate that the gamma-IFN response is transcriptionally activated but then subsequently subdued by feedback from the GAIT PTRO, thus optimizing a chronic inflammatory response and preventing an over-reaction of the immune system (Mukhopadhyay, Ray et al. 2008; Vyas, Chaudhuri et al. 2009).

(iii) In addition to most RBPs being associated with multiple mRNAs, it is has also been shown that the vast majority of mRNAs in yeast associate with multiple RBPs (Hogan, Riordan et al. 2008). This is another important aspect of global PTR- it is both cis and trans combinatorial, not unlike transcription. Trans combinatorial, as discussed above, describes the fact that most RBPs, as well as miRNAs, associate with multiple mRNA targets (Bartel and Chen 2004; Keene 2007). The cis-combinatorial aspect of coordination of PTR describes how most mRNAs have multiple binding and regulatory sites, both in the UTRs and coding sequence, which may potentially interact with trans-acting factors, another aspect of PTR demonstrated elegantly by Hogan et al. (Hogan, Riordan et al. 2008). Thus, combinations of different RBPs acting on a single mRNA and single RBPs interacting with multiple mRNAs are responsible for coordinating PTR.

(iv) PTR has been shown to occur in all species tested, and although the protein domains and sequence motifs responsible for interactions are well conserved, the identities of the proteins and mRNAs that contain these domains and motifs generally are not (Anantharaman, Koonin et al. 2002; Chan, Elemento et al. 2005; Gerber, Luschnig

et al. 2006; Galgano, Forrer et al. 2008; Morris, Mukherjee et al. 2008). This is another insight gained through ribonomics investigation; PTR evolves in a modular fashion, with the interacting portions of the RBPs and mRNAs being conserved but the identities of mRNAs that are coordinated and the consequences of this regulation differing across species. Thus, evolutionary re-wiring has occurred, during which selective pressures led to the conservation of RNA-binding domains and the bound sequence motifs, but allowed these interactions to be used in different functional settings that were best suited for survival (Mesarovic, Sreenath et al. 2004; Chan, Elemento et al. 2005; Gerber, Luschnig et al. 2006; Keene 2007; Halbeisen, Galgano et al. 2008). An excellent example of this re-wiring is described later in this study for the case of the PUF (for Pumilio/FBF) proteins, sequence specific RBPs that exist throughout the eukaryotic lineage (Wickens, Bernstein et al. 2002). The PUF RNA-binding domain, or PUM-HD, is extremely well conserved between yeast, fly and human PUF proteins (Wickens, Bernstein et al. 2002; Spassov and Jurecic 2003), as is the sequence motif to which this domain binds in global targets (Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006; Galgano, Forrer et al. 2008; Morris, Mukherjee et al. 2008). However, ribonomic analysis of PUF proteins from all of these species revealed that there was little conservation of the identities of the associated mRNAs or functions in which these mRNAs were involved, showing how evolutionary re-wiring of PTR occurs in a modular fashion (Gerber, Luschnig et al. 2006; Galgano, Forrer et al. 2008; Morris, Mukherjee et al. 2008).

(v) Another aspect of PTR previously revealed through ribonomic analysis was confirmed through studies of PUF proteins; RBPs tend to regulate other regulatory proteins, such as other RBPs and transcription factors (Mesarovic, Sreenath et al. 2004; Penalva, Burdick et al. 2004; Gerber, Luschnig et al. 2006; Pullmann, Kim et al. 2007; Morris, Mukherjee et al. 2008). This concept, depicted by the dotted lines in Figure 1, is referred to as the "regulators of regulators" concept, and was most elegantly demonstrated in a study by Gorospe and colleagues (Pullmann, Kim et al. 2007). This study revealed that a number of RBPs that bind to AU-rich elements in 3'UTRs of mRNAs (ARE-RBPs) can bind to and regulate mRNAs encoding other ARE-RBPs, forming a complex pattern of PTR (Pullmann, Kim et al. 2007). In fact, given that so many RBPs target mRNAs encoding other RBPs, it is likely that the PTR environment can be extremely robust by being self-sustaining and causing changes in gene expression independent of transcriptional input, at least for limited periods of time (Keene 2007; Mansfield and Keene 2009).

(vi) The final insight that will be discussed here is that the ribonome is highly dynamic; it responds to environmental signals by altering RNP contents, both protein and RNA, thus altering the gene expression program in a cell. A dramatic visual example of the dynamics of the posttranscriptional environment can be seen in the case of stress granules. Stress granules are large, cytoplasmic aggregates that contain various RBPs and mRNAs and form upon response to various types of stress (Anderson and

Kedersha 2006). Stress granules form within minutes of administration of stress, and will also begin to dissociate within minutes of stress being removed (Anderson and Kedersha 2008), demonstrating the rapid adaptability of PTR. Unfortunately, it has not been possible to isolate stress granules in order to analyze their RNA contents globally. An example of specific RNA dynamics of PTR was demonstrated by early RIP-Chip experiments of HuB, mentioned previously (Tenenbaum, Carson et al. 2000). This study showed that a subset of functionally related mRNA species associated with HuB changed in a coordinated manner during neuronal differentiation.

1.3 PUF Family RNA-Binding Proteins

PUF family RNA Binding Proteins (RBPs) are among the best characterized regulators of post-transcriptional gene expression in non-mammalian eukaryotes. Named for the founding members Pumilio (Lehmann and Nusslein-Volhard 1987)and FBF (Zhang, Gallegos et al. 1997), PUF proteins are represented throughout the eukaryotic lineage (Spassov and Jurecic 2003). The common feature of PUF proteins is the PUF Homology Domain (PUF-HD), an RNA binding domain typically consisting of eight imperfect repeats of a 32 amino acid sequence (Zamore, Williamson et al. 1997). The overall sequence of PUF proteins from different species is not highly similar outside of the PUF-HD, although the PUF-HD is incredibly well conserved (Spassov and Jurecic 2002). This extreme conservation of the PUF-HD suggests that post-transcriptional regulation of gene expression by PUF proteins has remained important throughout

evolution. Although there is a growing body of knowledge concerning PUF proteins in non-mammalian model organisms, at the outset of this study relatively little was known about the functions or mRNA targets of PUF proteins in mammals

Genetic analyses have revealed diverse functions of PUF proteins such as embryo patterning in Drosophila (Lehmann and Nusslein-Volhard 1987), germ line establishment and maintenance in Drosophila (Lin and Spradling 1997; Forbes and Lehmann 1998; Parisi and Lin 2000) and C. elegans (Zhang, Gallegos et al. 1997; Kraemer, Crittenden et al. 1999; Crittenden, Eckmann et al. 2003; Bachorik and Kimble 2005; Walser, Battu et al. 2006), and mitochondrial function in S. cerevisiae (Garcia-Rodriguez, Gay et al. 2007). PUF proteins have been found to function as repressors of gene expression through both repression of translation and enhancement of decay of target mRNAs (Wharton, Sonoda et al. 1998; Olivas and Parker 2000; Foat, Houshmandi et al. 2005; Goldstrohm, Hook et al. 2006; Goldstrohm, Seay et al. 2007). Crystal structure analysis of the PUF-HD from Drosophila Pumilio revealed that it forms a crescent shape (Edwards, Pyle et al. 2001), with protein-RNA interactions occurring on the inner concave surface and protein-protein interactions on the outer surface.

Genome-wide target identification of five S. cerevisiae PUF proteins, Puf 1-5 (Gerber, Herschlag et al. 2004), and the single Drosophila PUF protein, Pumilio (Gerber, Luschnig et al. 2006), demonstrated that each protein binds to a specific group of functionally related mRNAs distinct from those mRNAs bound by any other PUF

protein (except for Puf1 and Puf2, which bind overlapping sets of mRNAs). Puf3 was found to bind almost exclusively to mRNAs of nuclear-encoded mitochondrial proteins, and this binding was later found to have functional consequences when it was shown that Puf3 regulates stability of target messages in a condition-specific manner (Foat, Houshmandi et al. 2005) and regulates mitochondrial biogenesis and motility in S. cerevisiae (Garcia-Rodriguez, Gay et al. 2007). These experiments represent a compelling example of the post-transcriptional RNA operon/regulon model (Keene and Tenenbaum 2002; Keene 2007), demonstrating a mechanism through which expression of functionally related genes can be coordinately regulated at the level of the mRNA. Targets of Drosophila Pumilio also contained potential RNA regulons, most notably the vacuolar-type ATPase and the embryo-patterning cascade, which Pumilio mutants are known to disrupt (Gerber, Luschnig et al. 2006). These studies also demonstrated for the first time that while the cis-trans interactions between PUF proteins and target mRNAs are similar, target messages of PUF proteins are not conserved through evolution, at least from S. cerevisiae to Drosophila (Gerber, Luschnig et al. 2006).

At the beginning of this study, mammalian genomes were known two contain two PUF genes, Pum1 and Pum2 (Spassov and Jurecic 2002), both of which have been studied to a limited extent. The Pum2 gene has been knocked out in mouse, but the only obvious phenotype was smaller testis size with no effect on fertility (Xu, Chang et al. 2007). Several potential human Pum2 mRNA targets have been discovered by various

groups (White, Moore-Jarrett et al. 2001; Fox, Urano et al. 2005; Spik, Oczkowski et al. 2006; Lee, Hook et al. 2007); however the in-vivo genome-wide targets of the protein had not been identified at the outset of this study. Expression of reporter constructs containing Pum2 target 3'UTRs was shown to be repressed by Pum2 overexpression, but the mechanism of repression was not determined (Lee, Hook et al. 2007). Rat Pum2 was found to localize to stress granules in hippocampal neurons, with Pum2 knockdown interfering with stress granule formation and Pum2 overexpression inducing aggregates that co-stained with stress granule markers (Vessey, Vaccani et al. 2006). Even less is known about Pum1. The human Pum1 RNA-binding domain can bind to the Nanos Response Element, an mRNA sequence bound by Drosophila Pumilio, and has been found by crystal structure to interact with RNA in a very modular fashion, with each of the eight PUF repeats directly contacting a single RNA base (54) (Wang, McLachlan et al. 2002). Recombinant Pum1 was shown in vitro to interact with CNOT8 protein, a member of the CCR4-NOT deadenylase complex (Goldstrohm, Seay et al. 2007), suggesting that enhancement of target mRNA deadenylation and decay may be a conserved mechanism of PUF protein function (Wickens, Bernstein et al. 2002).

Recently, two more potential PUF genes were identified in mouse and human genomes, and one of the genes, PufA, has an orthologue that was characterized in zebrafish (Kuo, Wang et al. 2009). This orthologue is expressed in various types of stem cells in zebrafish and was found to function in primordial germ cell migration, further

supporting the ancestral role of PUF proteins in stem cell maintenance. PufA protein appears to bind RNA through a non-canonical PUF domain, with two of the PUF repeats being replaced by similar, yet distinct, protein sequences. These new sequences are thought to bind RNA similarly to PUF repeats, preserving the eight nucleotide recognition site (Kuo, Wang et al. 2009). The other newly identified PUF protein has not yet been characterized.

Due to the lack of knowledge regarding both targeting and function of mammalian PUF proteins, this study was undertaken to identify genome-wide mRNAs associated with Pum1 protein, along with mechanisms by which Pum1 regulates these mRNAs. By determining the genome-wide targets of a PUF protein from a third species, human, we gained insight as to how use of a highly conserved RNA-binding domain and cognate binding sequence has changed throughout evolution by regulating different sets of functionally related mRNAs. Identification of potential posttranscriptional RNA operons support the proposed ancestral function of PUF proteins in stem cell self-renewal, while results from mRNA decay and poly(A) tail length assays also support the hypothesis that enhancement of deadenylation is a conserved mechanism through which PUF family proteins repress their target mRNAs.

2. Identification of Pum1 associated mRNAs

2.1 Rationale

PUF family proteins likely exist in all eukaryotes, implicating them in important biological processes (Wickens, Bernstein et al. 2002). In fact maintenance of self-renewal of stem cells if the proposed ancestral function of PUF proteins (Wickens, Bernstein et al. 2002). When we started this study no in vivo mRNA targets of human Pum1 were known, although the protein had been shown in vitro to bind to a portion of the Drosophila *hunchback* mRNA known as the Nanos Response Element (Wang, McLachlan et al. 2002). The mRNA targets of RBPs often reflect the functions of those proteins (Morris, Mukherjee et al. 2009), thus we hypothesized that identification of mRNAs associated with human Pum1 may elucidate the functions of this protein in humans and other mammals.

RIP-Chip methodology developed in the Keene lab has proven highly effective for identifying mRNA targets of a variety of RBPs in a number of species (see reference (Morris, Mukherjee et al. 2009) for an extensive table). Thus, we employed this methodology in order to determine mRNA targets of human Pum1. These experiments, while quite informative, also served as a basis for a more detailed and systematic genome-wide target identification and ribonomic analysis of Pum1.

2.2 Results

2.2.1 Pum1 protein can be recovered from various cell types

We first tested the ability of the polyclonal anti-Pum1 antibody, BL289G, to immunoprecipitate (IP) Pum1 from a variety of cell types. We confirmed by Western blot and IP Western in Figure 3 that Pum1 protein was expressed and could be IPed from HeLa, LX-1, and Jurkat cells (not shown) using conditions optimized to recover entire RNPs, including both RNA and protein components (Tenenbaum, Carson et al. 2000).

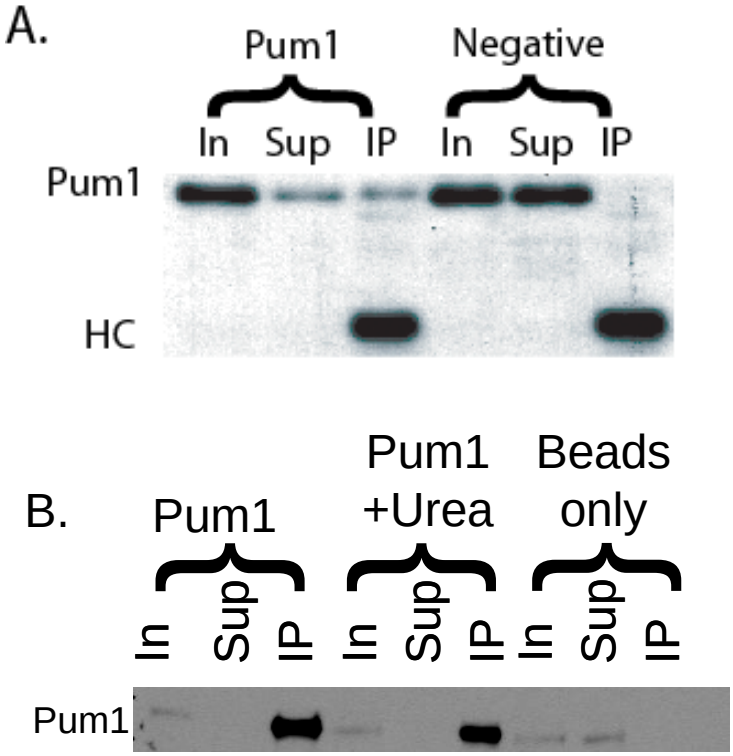


Figure 3: Pum1 immunoprecipitation. IP of Pum1 from HeLa (A) and LX-1 (B) cells. Frame (B) also confirms the success of the IP after supplementing 1M urea in three of the seven post IP washes. In=Input, Sup=Supernatant, IP=Pellet, HC=Antibody heavy chain

2.2.2 Pum1 is associated with Cyclin B1 mRNA

As Puf proteins from Drosophila and Xenopus bind to Cyclin B mRNA (Nakahata, Katsu et al. 2001; Gerber, Luschnig et al. 2006; Kadyrova, Habara et al. 2007), we hypothesized that this mRNA may also be bound by human Pum1. By extracting RNA from Pum1 IPs and comparing it to RNA extracted from control IPs and total RNA via RT-PCR, we confirmed that Cyclin B1 mRNA is indeed bound by Pum1 protein in all cell types tested (Figure 4).

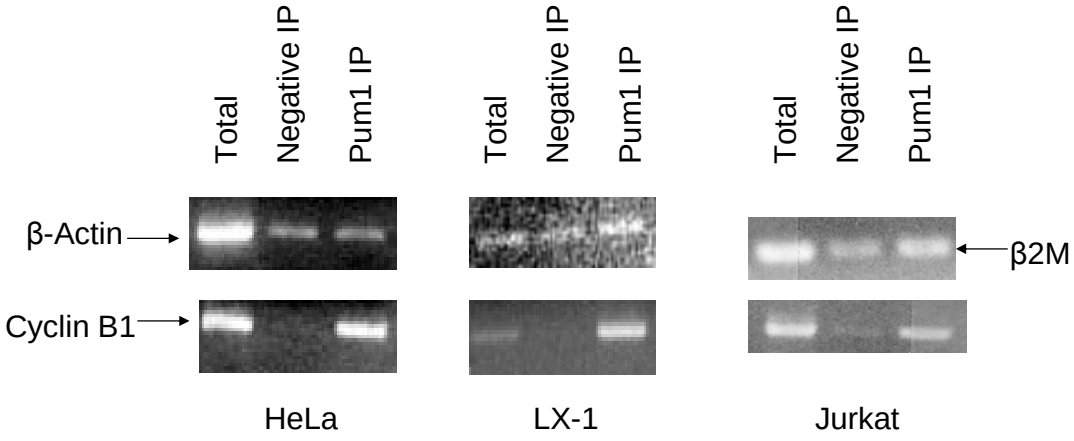


Figure 4: Pum1 binds Cyclin B1. RNA was extracted from input lysate (total), Pum1 IP pellet, and negative IP pellet and assayed by RT-PCR, demonstrating that Cyclin B1 mRNA is associated with Pum1 while β-actin or β-2-microglobulin mRNA is not.

2.2.3 Preliminary Pum1 RIP-Chip from LX-1 and Jurkat cells

In order to determine other potential targets of Pum1 we performed pilot RIP-Chip experiments in LX-1 and Jurkat cells (Table 1). Using cutoffs described in Methods, we determined that in LX-1 cells Pum1 bound to 4467 probes of 9571 total expressed, with the negative IP binding to 761. This large proportion of bound mRNAs indicated

that this IP had a high false positive rate, thus the data were not used as a set for further downstream analysis. Initial experiments from Jurkat cells revealed 1421 probes bound by Pum1 of 11559 expressed in total, with 583 present in the control IP. This proportion seemed more likely to represent actual Pum1 targets, thus analysis of this dataset was pursued further. Probes were collapsed into genes with unique UniGene IDs, resulting in a set of 513 mRNAs bound by Pum1 after correcting for the control IP, out of 7027 mRNAs expressed in total. Gene Ontology analysis of these targets using the program GOTM (Zhang, Kirov et al. 2005) revealed a significant enrichment for the category "Regulation of Progression through Cell Cycle." As Puf proteins have proposed ancestral roles in maintenance of stem cells and the differentiation/proliferation decision (Wickens, Bernstein et al. 2002), this category seemed logical. A number of these cell cycle related targets were confirmed by semi-quantitative RT-PCR and quantitative RT-PCR (RT-qPCR), as shown in Figures 5 and 6. Further attempts at Pum1 RIP-Chip experiments from Jurkat cells failed, and at this point we decided to switch to HeLa cells for further analysis of Pum1.

Table 1: Pilot RIP-Chip Results from LX-1 (A) and Jurkat (B) cell lines. Presence of a spot was determined by signal two-fold above local background. For Jurkat experiment, Pum1 IP correction and Gene Ontology enrichment were determined as described in Methods.

A.

Category	Number of probes detected
LX-1 Total expressed	9571
LX-1 Pum1 IP	4467
LX-1 Negative IP	761
Jurkat Total expressed	11559
Jurkat Pum1 IP	1421
Jurkat Negative IP	583

B.

Category	Number of unique genes
Jurkat total expressed	7027
Jurkat corrected Pum1 IP*	513

*Significant enrichment for Gene Ontology category *Regulation of Progression through Cell Cycle, P=0.0007*

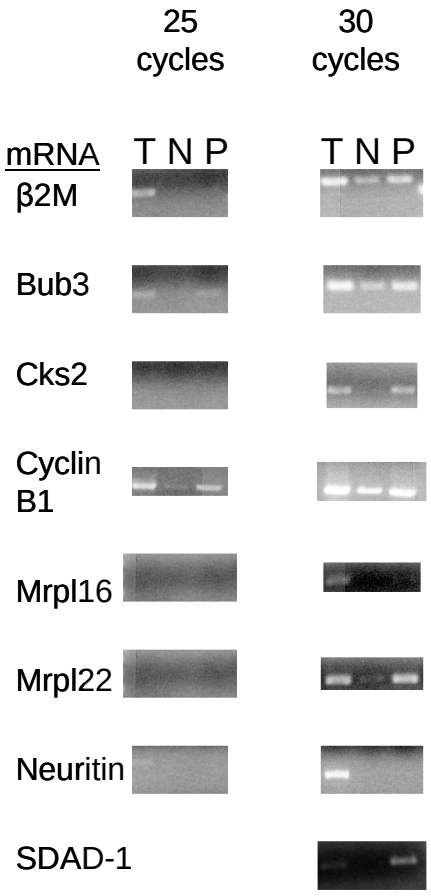


Figure 5: Semi-quantitative RT-PCR confirmation of Pum1 targets in Jurkat cells. All mRNAs shown except β2M and neuritin were predicted to be Pum1 targets based on RIP-Chip. T=Total RNA, N=Negative IP, P=Pum1 IP

It was decided that the ribonomic analysis of Pum1 would continue in HeLa cells for a variety of reasons. Pum1 protein expression is relatively higher in HeLa cells than Jurkat (Figure 7(A)), indicating that it may be easier to isolate more Pum1 protein and bound mRNA from HeLa cells. This was confirmed via RT-qPCR experiments which demonstrated a much greater abundance of Pum1 target mRNA in IPs performed from HeLa cells (Figure 7(C)). This greater abundance of Pum1 protein and increased ability

to isolate associated mRNA indicated that experiments in HeLa cells were likely to lead to more robust RIP-Chip results. In addition, HeLa cells express relatively low levels of the only other human Puf protein, Pum2 (Figure 7(B)). This is advantageous because the two proteins are very similar, thus any downstream experiments involving perturbation of Pum1 protein would be less likely to be complicated by compensation from Pum2. HeLa cells are more amenable to various molecular biology techniques, including immunofluorescence, transfection and expression of reporter constructs, and si-RNA mediated protein knockdown, increasing the potential "toolbox" for downstream experiments to further explore functions of Pum1. HeLa cells are also amenable to cell cycle synchronization (Firket and Mahieu 1967), and since preliminary results indicated that Pum1 may have a role in cell cycle, this was also advantageous.

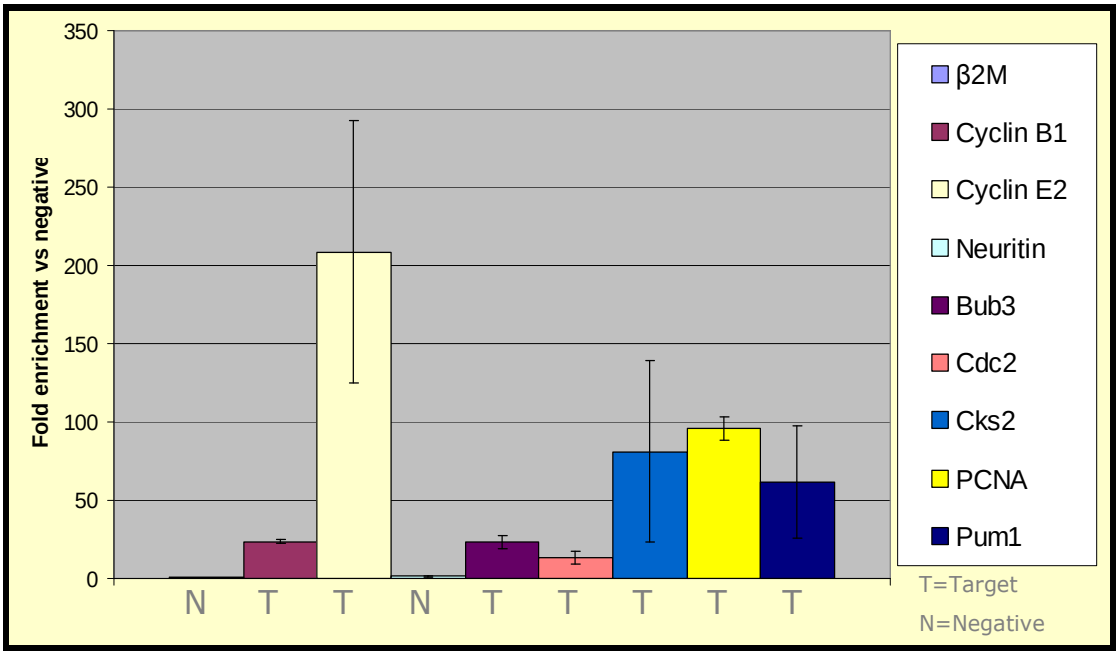
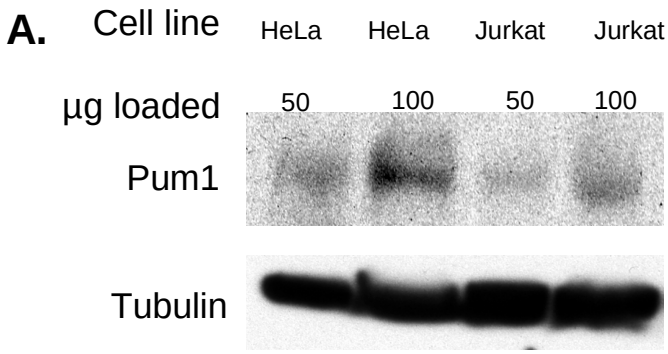
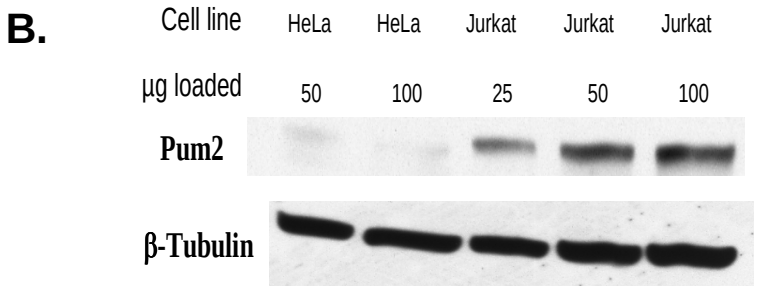


Figure 6: Quantitative RT-PCR confirmation of Pum1 targets in Jurkat cells. All mRNAs shown except β2M and neuritin were predicted to be Pum1 targets based on RIP-Chip. The non-target mRNA β2M was used for normalization. Bars represent SEM of 3 replicates.





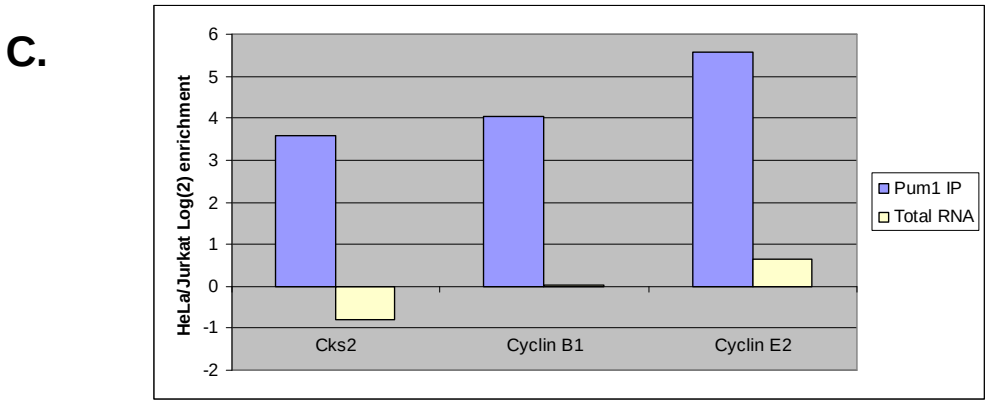


Figure 7: Comparison of HeLa and Jurkat cell lines. (A) and (B) are Western blots showing relative protein expression of Pum1 and Pum2, while quantitative RT-PCR in (C) demonstrates abundance of Pum1 target mRNAs in Pum1 IPs and total RNA from HeLa versus Jurkat cell lines.

2.3 Methods

2.3.1 Cell Culture

HeLa S3 cells were grown in Hams F12 supplemented with 10% FBS. Jurkat and LX-1 cells were grown in RPMI 1640 media supplemented with 10% FBS. All cells were grown in humidified incubators at 37 degrees and 5% CO2.

2.3.2 Immunoprecipitation

IP reactions were performed as described previously (Keene, Komisarow et al. 2006). Briefly, 10ug of anti-Pum1 antibody (goat polyclonal, Bethyl Labs) was incubated overnight with protein G-agarose beads. The beads were washed, then buffer and cell lysate were added and the reactions tumbled for four hours at four degrees. After this incubation, the beads were thoroughly washed again then either boiled in 2x Lamelli buffer for IP-Western experiments or had 1ml TRIzol added and RNA extracted for microarray and RT-PCR experiments. Identical IPs performed with beads pre-coated with pre-immune goat serum were used as a negative control.

2.3.3 Western Blots

Protein samples were loaded onto 4-20% Tris-HCl PAGE gels. After electrophoresis, protein was transferred to nitrocellulose membranes, and then these membranes were blocked and probed with the same anti-Pum1 antibody used for immunoprecipitations. Pum2 antibody was from Bethyl Labs. Visualization was performed with HRP-linked secondary antibody and ECL detection.

2.3.4 RNA extraction

TRIzol was used for all RNA extraction according to the manufacturer's protocol.

2.3.5 RT-PCR

Reverse transcription was performed with the iScript cDNA synthesis kit from Bio-Rad according to manufacturer protocols, using a combination oligo d(T) and random hexamers for priming. End-point PCR was carried out in the linear range (25 or 30 cycles) and products were resolved on 1% agarose gels and stained with ethidium bromide. qPCR was performed using a Roche Lightcycler with SYBR Green detection (Invitrogen) and delta delta CT analysis method, using either β-2-microglobulin (B2M) or GAPDH for normalization.

2.3.6 Microarray Analysis

Arrays were printed at the Duke Microarray Facility using the Genomics Solutions OmniGrid 300 Arrayer. The arrays contained Human Operon v3.0.2 oligo set (Oligo Source) that consists of 34,602 unique optimized 70-mers. RNA quality was ascertained using an Agilent 2100 bioanalyzer (Agilent technologies). GPR files were filtered for probes containing signal greater than twofold over local background, then all probes still present in the positive IP that were in the negative IP but whose signal/background ratio was not at least tenfold greater in the positive IP versus the negative IP were also removed. The resulting Pum1 IP and total RNA probe sets were collapsed into unique UniGene symbols using Genespring.

2.3.7 Gene Ontology Analysis

Gene Ontology Tree Machine (GOTM) was used to analyze GO enrichment of Pum1 targets from Jurkat cells, using total RNA as the background model (Zhang, Kirov et al. 2005).

2.4 Discussion

Pum1 has been previously reported to be expressed at the mRNA level in a wide variety of tissues (Spassov and Jurecic 2002), and we confirmed that it is also expressed in 3 somewhat diverse human tissue culture lines. We also confirmed that using RIP-Chip methodology we were able to recover Pum1 protein and associated mRNA from all 3 cell lines.

Although Cyclin B was previously reported to be a target of PUF proteins in other species (Nakahata, Katsu et al. 2001; Gerber, Luschnig et al. 2006; Kadyrova, Habara et al. 2007), this was the first demonstration that it is also a PUF protein target in humans, indicating that it is likely an ancestral target of PUF proteins and will be a target of at least one PUF protein in many species. This knowledge is advantageous because knowing an mRNA target of an RBP greatly assists in optimization of more global RIP-Chip type experiments, or any type of experiments where recovering an intact RNP is important. In fact, the discovery of Cyclin B1 as a human Pum1 target helped us to determine that we were indeed able to recover mRNA associated with Pum1 protein, increasingly the likelihood of an informative RIP-chip experiment.

3. Ribonomic analysis of human Pum1

3.1 Rationale

Analysis of genome wide targets of mRNA binding proteins has revealed that there is typically a functional relationship between the proteins encoded by those mRNAs (Morris, Mukherjee et al. 2009), forming posttranscriptional RNA operons (Keene and Tenenbaum 2002). An example of this is the aforementioned yeast PUF protein Puf3, which binds almost exclusively to mRNAs of nuclear encoded mitochondrial protein genes (Gerber, Herschlag et al. 2004) and regulates mitochondrial biogenesis and motility (Garcia-Rodriguez, Gay et al. 2007). Although often not as striking, examples of this type of functional relationship between mRNAs associated with an RBP pervade the literature, as was predicted by the PTRO model (Keene and Tenenbaum 2002).

Another common result predicted by the PTRO model is that there is often a motif to which the RBP binds in the 5′ or 3′ UTRs of the target mRNAs which can be identified through analysis of those target mRNAs as identified by RIP-Chip (Keene and Tenenbaum 2002). This type of motif has been found through ribonomics analysis of yeast and fly PUF proteins (Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006) and is thus likely to be identifiable in targets of human Pum1.

The human PUF protein Pum1 is widely expressed (Spassov and Jurecic 2002), and while our group had determined some mRNAs associated with Pum1, the genome-

wide targets had not been identified and analyzed in a systematic manner. We sought to use ribonomic analysis based on RIP-Chip to identify the genome-wide mRNA targets of human Pum1, biologically relevant associations between these targets, and a consensus binding site for Pum1 (Tenenbaum, Lager et al. 2002), as well as to compare mRNA targets of Pum1 to targets of PUF proteins from other species. These observations should in turn help to elucidate the functions of Pum1.

3.2 Results

3.2.1 Pum1 RIP-Chip from HeLa cells

We performed RIP-Chip in biological triplicate in HeLa cells using custom spotted cDNA microarrays. Each biological replicate consisted of a Pum1 IP sample, a negative control IP sample, and a total RNA sample, which were each hybridized to a separate microarray along with a common reference sample. Total RNA microarrays were used to identify the transcriptome of the cells from which IPs were performed, providing an accurate background for subsequent analyses. A probe was considered present in the transcriptome if the signal from the spot was at least two-fold greater than local background in all three total RNA or Pum1 IP microarrays. T-scores, based on the T-statistic, for Pum1 IP versus negative IP were calculated for all probes on the microarrays (Subramanian, Kuehn et al. 2007). A visual inspection of the T-score values of the probes (Figure 8, histogram) suggested two distributions; a background distribution and a Pum1 associated distribution.

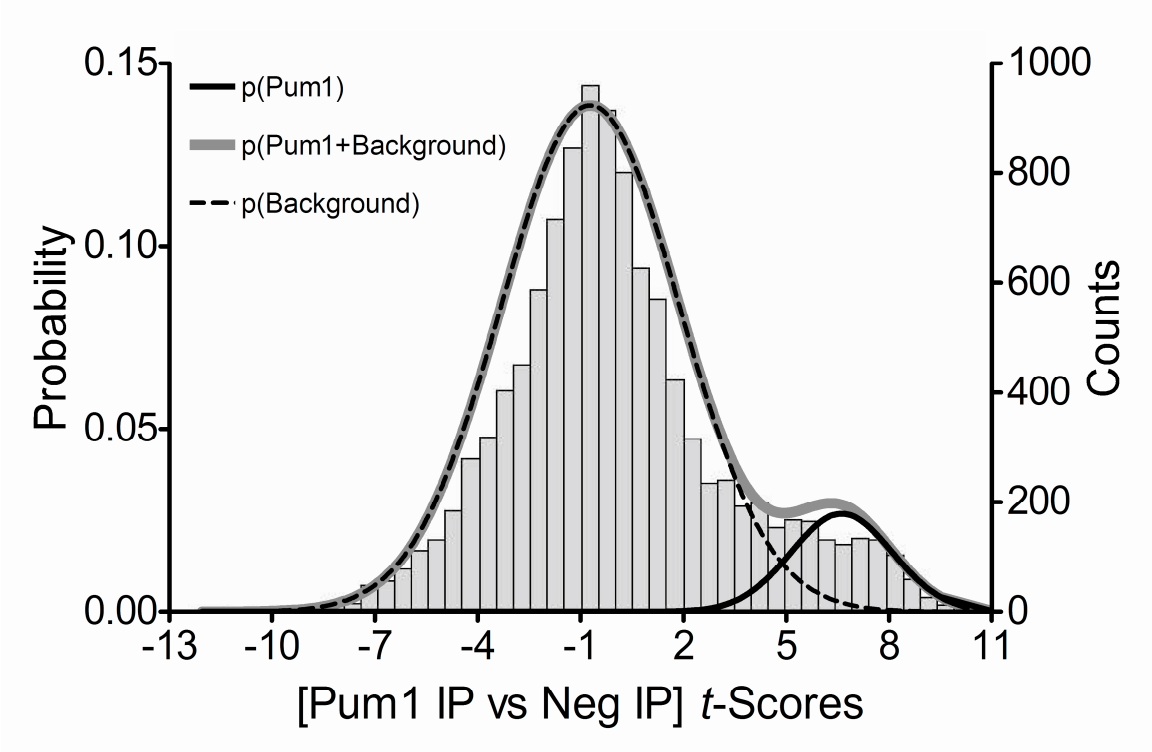


Figure 8: Pum1 RIP-Chip results. Distribution of Pum1 IP versus negative IP T-scores for 3 biological replicates of RIP-Chip are shown. The x-axis represents T-scores, the right y-axis represents number of probes and the left y-axis represents probability. Pum1-associated (solid black curve), background (dashed black curve), and the sum of Pum1-associated and background (gray curve) probability distributions are shown, as defined by Gaussian mixture modeling.

3.2.2 Gaussian Mixture Modeling of Pum1 RIP-Chip data

In order to objectively define Pum1 target mRNAs based on a distribution of T-scores we employed Gaussian mixture modeling (GMM). GMM uses probabilistic modeling to identify single Gaussian distributions in a population consisting of a mixture of multiple Gaussian distributions (Pearson 1894). GMM uses expectation maximization modeling which is prone to converging on a local optimum, therefore several iterations of mixture modeling were performed which initialized at different

values in the distribution, and the model with the greatest probability was then used to define Pum1-associated mRNAs. Both Pum1 and background distributions were defined as an equation relating T-scores to probability (Figure 8, curves), and using these equations we calculated the log of odds (LOD) of each probe being in the Pum1 distribution versus the background distribution. Those probes with a greater likelihood of being in the Pum1 distribution (LOD>0) were considered Pum1 associated mRNAs (targets). While one might expect that this cutoff would result in a high false positive rate, downstream analysis proved that a cutoff of LOD>0 was appropriate. The list of probes, T-scores, and LOD scores is available as Supplementary Table 1 in reference (Morris, Mukherjee et al. 2008). Use of GMM and creation of LOD scores allowed us to objectively define probabilities of probes being in the Pum1 distribution versus the background distribution, and thus allowed for a less arbitrary determination of probes that could be considered Pum1 associated. Probes were collapsed into unique genes, and of the 6539 unique genes represented in total, 726 (11.1%) were considered Pum1 targets on this basis.

3.2.3 Confirmation of Pum1 targets

We used RT-qPCR to confirm select targets by measuring their enrichment in the Pum1 IP versus either the negative IP or total RNA, using the non-target message β2M for normalization and GAPDH mRNA as a negative control. qPCR analysis confirmed levels of enrichment up to 240 fold for target messages in the Pum1 IP, with levels of

enrichment being similar when compared to either the negative IP or total RNA (Figure 9). Target messages confirmed by RT-qPCR represented a range of LOD scores greater than 0, thus the confirmation of these targets serves as partial validation of the LOD>0 cutoff when defining Pum1 associated mRNAs, as noted above.

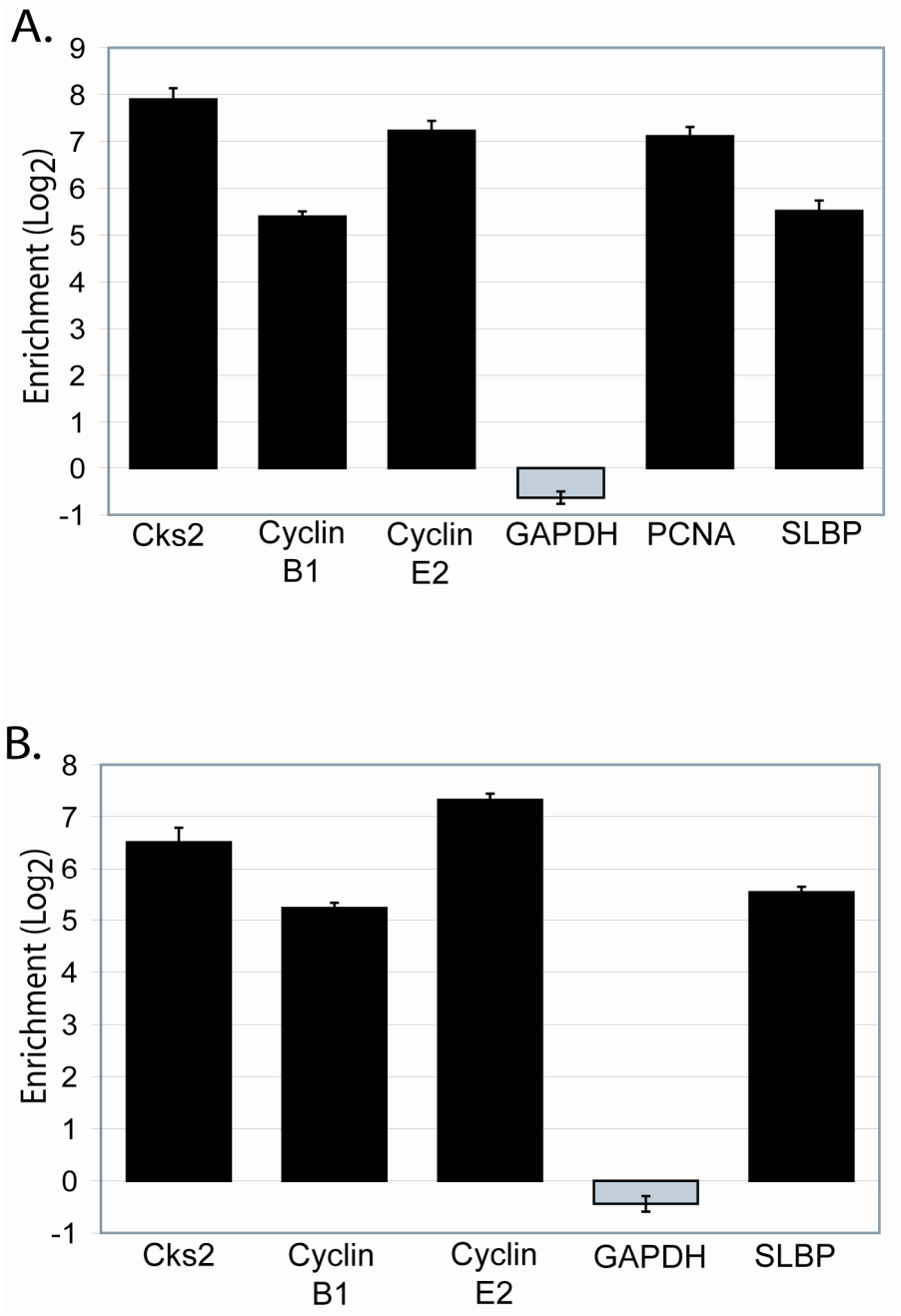


Figure 9: RT-qPCR confirmation of Pum1 associated mRNAs. Log2 enrichment versus either total RNA (A) or negative IP (B), normalized to β-2-microglobulin, is shown. Select target messages are represented by black bars and non-targets by grey bars. Error bars represent SEM of 3 biological replicates.

3.2.4 Pum1 associated mRNAs share functional relationships

After defining mRNAs associated with the Pum1 RBP, we proceeded to determine if the proteins encoded by Pum1 target mRNAs were functionally related, as is predicted by the RNA operon model (Keene and Tenenbaum 2002; Keene 2007). We analyzed Pum1 target genes using two web-based programs, Panther (Mi, Guo et al. 2007) and WebGestalt (Zhang, Kirov et al. 2005), which search gene lists for significant enrichment in gene ontology (GO) categories and other functional groupings. We also analyzed Pum1 targets using Gene Set Enrichment Analysis (GSEA), which compares a rank-ordered gene list of interest to other gene sets in the GSEA Molecular Signatures Database to discover significant correlations between sets of genes (Subramanian, Kuehn et al. 2007). Results from WebGestalt and Panther, as seen in Table 2, were very similar. The Panther program searches for both positively and negatively enriched categories and applies a strict Bonferroni correction for multiple testing when calculating significance (Mi, Guo et al. 2007), thus fewer categories were found to be enriched in the Panther analysis and the statistical significance of that enrichment was lower. WebGestalt makes no correction for multiple testing, thus the results obtained by this method, while having lower p-values, may be less biologically significant. However, results obtained using either WebGestalt or Panther were in agreement. One of the more striking results of both the WebGestalt and Panther analyses was the large number of target genes and extreme significance of enrichment of GO categories

involved in transcription and regulation of transcription. Another noteworthy result was the enrichment of mRNAs representing genes involved in regulation of cell cycle and cell proliferation and differentiation, a result consistent with proposed ancestral functions of PUF proteins in stem cell biology (Wickens, Bernstein et al. 2002). GSEA analysis of the Pum1 IP rank-ordered gene list created from LOD scores revealed a high degree of correlation with various gene sets, including gene sets whose mRNAs levels were found to decrease after UVC and UVB exposure, whose mRNAs increased after CMV infections, a gene set consisting of HOX genes, and gene sets related to cell cycle (Table 2). This result provides further support for a role for Pum1 protein in stem cell function, as well as a role for Pum1 in response to stress (see below).

Table 2: Functional enrichment of Pum1 associated mRNAs. "No. of genes" represents number of Pum1 associated mRNAs found in this category.

Analysis method and GO category (positively or negatively enriched)	No. of genes	*P* value
Panther		
Positively enriched		
mRNA transcription	100	9.66E-04
mRNA transcription regulation	71	9.40E-03
Transcription factor	103	3.02E-05
Other transcription factor	32	4.42E-02
Nucleoside, nucleotide, and nucleic acid metabolism	173	2.57E-02
Cell proliferation and differentiation	51	1.76E-02
Negatively enriched		
Oxidoreductase	5	9.90E-07
Reductase	0	1.15E-02
Dehydrogenase	2	2.85E-02
Electron transport	2	1.32E-02
WebGestalt		
Positively enriched		
Regulation of biological processes	220	1.17E-09
Regulation of transcription	138	5.48E-10
Transcription factor activity	67	1.10E-09
Regulation of cell cycle	42	2.71E-03
Ubiquitin cycle	30	5.80E-03
Wnt receptor signaling pathway	11	1.95E-03
Small GTPase-mediated signal transduction	32	3.75E-04
Cell communication	128	2.13E-04
GSEA[a]		
Positively enriched		
UVC high all DN	117	<0.001
UVB NHEK1 DN	66	<0.001
CMV/HCMV time course all up	76	<0.001
HOX genes	10	<0.001
Cell cycle	28	<0.001

[a] For GSEA results, the numbers of genes reported represent core enrichment. The GO category names for GSEA are those used in the analysis software.

Within the set of Pum1 target genes are several specific functional relationships that represent putative RNA operons/regulons, such as that between Cyclin B1, Cdc2, p21 and Wee1. The Cyclin B1-Cdc2 complex is a key regulator of the G2/M transition of mitotic cell cycle, with p21 and Wee1 acting as negative regulators of Cdc2 (www.biocarta.com). Although it may seem counter-intuitive that Pum1 would regulate expression of both one protein and a second protein that negatively regulates the first protein, this situation has been seen in C. elegans, where a PUF protein represses a MAP kinase and a gene that inactivates the same MAP kinase, thereby ensuring continued repression of the MAP kinase gene after PUF mediated repression of both proteins is

relieved (Lee, Hook et al. 2007). PCNA, GSK3β, p21, and p27 form another cell cycle related functional group, as all of these proteins function as inhibitors of Cyclin D (www.biocarta.com). One of the most striking potential Pum1 regulons is that of the E2F transcription factors: four of the five E2Fs that were represented in the total mRNA population were found to be Pum1 targets (E2Fs 3-6 are targets, E2F1 not a target), showing an overlap of two highly enriched categories in Pum1 targets; cell cycle/proliferation and transcription (DeGregori and Johnson 2006). A large number of RNA processing and RNA binding protein genes were also found to be Pum1 targets, among them the histone Stem Loop Binding Protein (SLBP), DICER, Pum1 itself, and the other human PUF protein, Pum2.

3.2.5 Identification of the Pum1 Untranslated Sequence Element for Regulation (USER)

Previous analyses of PUF protein target mRNAs revealed a consensus sequence present in the 3'untranslated regions (UTRs) of target messages (Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006). We used the motif finding program MEME (Bailey, Williams et al. 2006) to search for a consensus sequence in the 3' UTRs of Pum1 targets. The 3'UTR sequences of the top 100 Pum1 associated mRNAs, by LOD score, were used as a training set for MEME analysis, resulting in discovery of the consensus sequence shown in the inset in Figure 5. Contained in the Pum1 target consensus sequence is the eight nucleotide core sequence UGUAHAUA, which has been shown by X-ray crystallographic analysis to be directly bound by the Pum1 PUF-HD (Wang, McLachlan

et al. 2002). We searched for the occurrence of this core sequence in the 3′ UTRs of the Pum1 targets not used in the MEME training set and found it occurred at least once in 46.5% of Pum1 target 3′UTRs but only in 13.5% of total mRNA 3′UTRs. To determine the likelihood of this enrichment occurring by chance, we created 50,000 sets of random 3′ UTRs and determined the frequency of 3′ UTRs in each set that contained the core consensus sequence at least once. The random sets of 3′ UTRs were chosen from the mRNAs expressed in HeLa cells, and each set contained the same number of 3′UTRs as the Pum1 associated set. The occurrence of the Pum1 core consensus sequence was determined for each of these sets, resulting in a distribution with a mean of 13.5% and a maximum of 20.1% (Figure 5). This represents an extremely significant enrichment of the core Pum1 consensus sequence in Pum1 targets, $p<2x10\text{-}5$, even after excluding those 3′UTRs used as the training set. We also used Fisher's exact test to determine the significance of enrichment of this sequence in Pum1 targets and calculated it to be: p=1.99E-60. As elements of this eight nucleotide core sequence have been shown to be important for target mRNA regulation by PUF proteins (Murata and Wharton 1995; Crittenden, Bernstein et al. 2002; Jackson, Houshmandi et al. 2004; Bernstein, Hook et al. 2005; Goldstrohm, Hook et al. 2006; Lee, Hook et al. 2007), we will henceforth refer to this sequence as the Pum1 USER (Untranslated Sequence Element for Regulation) (Keene and Tenenbaum 2002).

Although it may be unexpected to find that only about half of Pum1 target messages contained the Pum1 USER, there are a number of likely explanations for this outcome. The RIP-Chip method is optimized to isolate entire RNPs (Tenenbaum, Carson et al. 2000; Keene, Komisarow et al. 2006), and thus some of those messages associated with Pum1 RNPs are not expected be directly bound by Pum1. The search for the Pum1 USER was also based on a simple string (UGUAHAUA) rather than a more descriptive and flexible position-specific weight matrix. Because the position-specific weight matrix more accurately describes flexibility in the consensus sequence, it could also identify sequences in 3'UTRs to which Pum1 binds with almost as high affinity as the consensus sequence, which were not recognized by a simple string search. Finally, only the 3' UTRs of mRNAs were searched for the Pum1 USER, thus any USERs present in the 5' UTRs or coding sequences would not be identified.

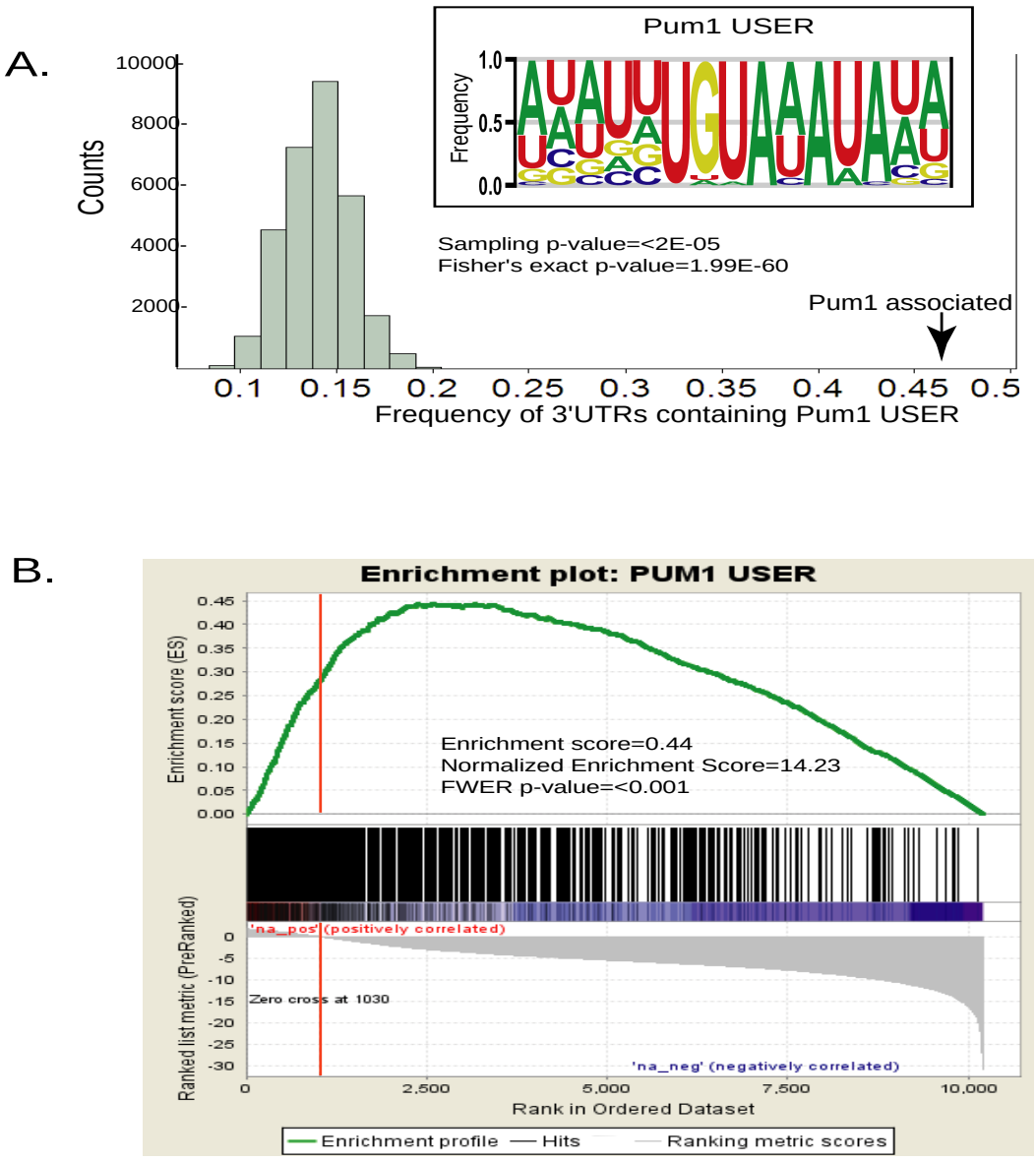


Figure 10: Identification and enrichment of the Pum1 USER. A. The calculated frequency of 3′ UTRs containing the Pum1 Untranslated Sequence Element for Regulation (USER), UGUAHAUA, among Pum1 associated mRNAs compared with 50,000 randomly chosen sets of mRNAs. P-values derived from this sampling, and using Fisher's exact test, are shown. Random sets of mRNAs were derived from mRNAs present in the HeLa cell transcriptome and each set contained the same number of mRNAs as the Pum1 associated set. These sets are represented by the y-axis and the frequency of the Pum1 USER in these sets is represented on the x-axis. Inset shows the consensus Pum1 target sequence, including the eight nucleotide

Pum1 USER, as determined by MEME analysis. B. GSEA analysis of the enrichment of the Pum1 USER based on LOD scores. Shown is the running enrichment of Pum1 associated mRNAs containing the Pum1 USER as described in (Subramanian, Kuehn et al. 2007). Enrichment scores and the Familywise-Error Rate (FWER) p-value are show. The vertical red line in the graphs represents LOD=0.

In order to more thoroughly explore enrichment of the Pum1 USER in Pum1 associated messages, we also used GSEA analysis (Subramanian, Kuehn et al. 2007). This analysis used all genes with the Pum1 USER in their 3′ UTR as a gene set and calculated the running enrichment of genes containing the Pum1 USER in the Pum1 RIP-Chip data ordered by LOD score. As can be seen in Figure 5B, the peak of the running enrichment score occurs after LOD=0, demonstrating that our earlier cutoff of LOD>0 is valid for determining targets, and is likely conservative. The normalized enrichment score of 14.23 and FWER p-value of <0.001 for this analysis again demonstrates the extreme enrichment of this sequence in the 3′UTRs of Pum1 associated mRNAs.

3.2.6 Conservation of PUF protein target mRNAs and potential RNA Regulons

One of the main goals of this study was to compare Pum1 target genes to target genes of PUF proteins in other species in order to observe how the post-transcriptional regulatory adaptors of the PUF-HD and cognate binding sequence have been rewired throughout evolution. The four amino acids that directly contact RNA in each of the eight repeats of the PUF-HD (Wang, McLachlan et al. 2002) are completely conserved between S. cerevisiae Puf3, Drosophila Pumilio, and human Pum1, as is the eight nucleotide sequence to which each PUF HD likely binds (Gerber, Herschlag et al. 2004;

Gerber, Luschnig et al. 2006). Neither the RNA contacting amino acids or consensus binding sequence of any other yeast Puf protein show this level of conservation, thus only Puf3 was considered in this analysis. mRNA targets of Puf 3 and Pumilio have been previously determined using RIP-Chip, and it was found that there was little conservation of target genes between them (Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006). In order to determine if there is conservation of targets between Pum1 and Puf3 or Pumilio, we first determined how many Puf3 or Pumilio target genes have human orthologues that are expressed in HeLa cells, and then determined how many of those genes with human orthologues are also Pum1 targets (Figure 11(A)). Puf3 target genes have 89 human orthologues, and only 7 of these are also Pum1 target genes, representing no significant enrichment for Pum1 targets among Puf3 targets. Pumilio targets in the adult Drosophila ovary have 502 human orthologues, with 73 also being Pum1 target genes. This represents a slightly significant enrichment of Pum1 targets among Pumilio targets, with a sampling p-value of 0.0183 and a Fisher's exact p-value of 0.036. Statistical significance via sampling was determined by creating 10,000 sets of random genes containing the same number of genes as the Puf3 or Pumilio target genes with orthologues, then determining the proportion of those genes that were Pum1 targets. A possible caveat to this analysis is that each RIP-Chip experiment was performed in a different cell type and thus each PUF protein had the potential to

associate with different mRNAs in vivo. To address this issue we limited the universe of this analysis to only genes that were expressed in HeLa cells as defined by our study.

A.

	Genes with human orthologues	Pum1 target genes	Percent conserved target genes	P-value of enrichment (sampling)	P-value of enrichment (Fisher's exact)
Puf3 target genes	89	7	7.87%	0.83	0.79
Pumilio target genes	502	73	14.54%	0.018	0.036

B.

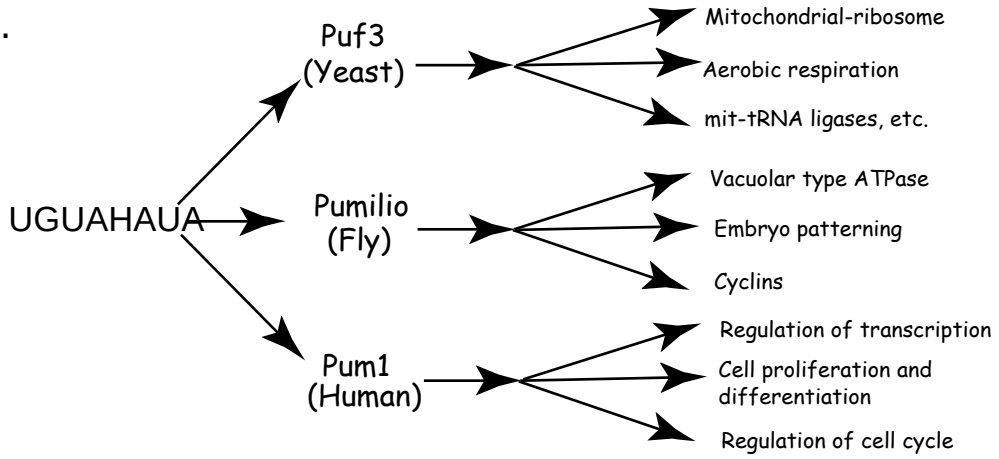


Figure 11: Evolutionary rewiring of the PUF domain. A. The number of S. cerevisiae Puf3 and Drosophila Pumilio target genes that have human orthologues, and the number of these that are also Pum1 targets in HeLa cells. The statistical significance of this conservation of targets is indicated. B. The motif UGUAHAUA is predominant in yeast Puf3, fly Pumilio, and human Pum1 associated mRNAs, while the proteins encoded by each PUF protein's target mRNAs are not conserved, consistent with the notion of evolutionary re-wiring of RNA regulon networks. Data on Puf3 and Pumilio targets were derived from (Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006).

In addition to identities of genes encoded by target mRNAs, we also considered whether functional relationships among target genes, or potential RNA regulons (Keene

and Tenenbaum 2002; Keene 2007), were conserved (Figure 11(B)). Target genes of Puf3 are almost entirely nuclear encoded mitochondrial proteins (Gerber, Herschlag et al. 2004), while those of Pumilio and Pum1 are not enriched for mitochondrial functions (Gerber, Luschnig et al. 2006). The potential RNA regulon of cell cycle regulation is not conserved between Pumilio and Pum1, although there are many cyclin mRNAs among the Pumilio targets (Gerber, Luschnig et al. 2006). Conversely, Pum1 targets are not enriched for components of the V-type ATPase or the embryo patterning cascade, functional groupings found to be enriched among Pumilio targets (Chan, Elemento et al. 2005; Gerber, Luschnig et al. 2006). Pum1 targets and Pumilio targets are both enriched for the GO categories of transcription factor and membrane-bound organelle (Gerber, Luschnig et al. 2006), showing that while most of the individual genes regulated by the proteins have changed, the types of genes regulated are similar. Combined with data regarding conservation of individual targets, this observation indicates that neither targets nor functions of PUF proteins have been conserved between S. cerevisiae and human, yet some targets and functions have been conserved between Drosophila and human. Thus, by keeping the PUF HD and cognate binding sequence as static modules but changing the sets of mRNAs that are regulated, evolution has been able to re-wire post-transcriptional regulation of gene expression networks, as previously suggested through experimental and computational analysis of evolutionary conservation of potential post-transcriptional regulatory elements (Mesarovic, Sreenath et al. 2004; Chan,

Elemento et al. 2005; Gerber, Luschnig et al. 2006; Keene 2007; Halbeisen, Galgano et al. 2008).

3.3 Methods

3.3.1 Cell culture

HeLa S3 cells, used for all experiments from this point forward except immunofluorescence, were grown in Hams F12 supplemented with 10% FBS.

3.3.2 Microarray Analysis of HeLa data

Arrays were printed at the Duke Microarray Facility using the Genomics Solutions OmniGrid 300 Arrayer. The arrays contained Human Operon v3.0.2 oligo set (Oligo Source) that consists of 34,602 unique optimized 70-mers. RNA quality was ascertained using an Agilent 2100 bioanalyzer (Agilent technologies). All microarray data were submitted to the GEO database and can be found under the accession ID GSE 11301. All arrays were subject to loess normalization within each array and scale normalization across all arrays using the Array Magic package in R (Buness, Huber et al. 2005). Replicate probes were collapsed to one probe corresponding to the median value of all the replicates. GSEA was used to calculate T-scores comparing the Pum1 IP to the NGS IP (Subramanian, Kuehn et al. 2007).

3.3.3 Gaussian mixture modeling

Gaussian mixture modeling was performed multiple times on the T-score distribution to estimate the mean, standard deviation and weight of each mixture using

the Mixtools package in R (Pearson 1894). Each iteration of mixture modeling initialized at different values in the distribution, and the parameters from the model with the highest likelihood were used to create LOD scores of Pum1 RNP-association by comparing the weighted probability density functions of the Pum1 associated versus background distributions. LOD scores greater than 0 have a higher probability of being in the Pum1-RNP associated distribution compared to background distribution, therefore a LOD score of 0 was used as a cut-off for determining Pum1 associated probes.

3.3.4 Functional enrichment

Pum1 associated and total expressed genes were loaded into either WebGestalt (Zhang, Kirov et al. 2005)or Panther (Mi, Guo et al. 2007) using appropriate gene identifiers. Pum1 target genes were compared to total expressed to determine functional enrichments. To calculate significance of enrichment, WebGestalt uses a hypergeometric test and Panther uses a modified Fisher test with a Bonferonni correction for multiple testing.

3.3.5 Motif finding

3′ UTRs of Pum1 target and total expressed genes were obtained from a local pipeline which uses information from PolyA_DB (Zhang, Hu et al. 2005) combined with other data to define 3′UTRs by genomic coordinates (Majoros and Ohler 2007), and sequences were obtained from the latest human genome build based on these

coordinates. MEME analysis (Bailey, Williams et al. 2006) was run locally using 3′ UTRs of the top 100 Pum1 associated mRNAs as a training set.

3.3.6 Orthologues

Orthologues of Puf3 and Pumilio target genes were determined using the online database mining tool Biomart (www.biomart.org).

3.4 Discussion

This study represented the first ribonomic analysis of a human PUF protein, and perhaps more importantly the first use of Gaussian mixture modeling to analyze RIP-Chip data. GMM allowed us to make a more sophisticated, probabilistic determination of which probes were Pum1 associated and which represented background. By calculating probabilities of association, this method also allows for more advanced downstream analysis of target mRNAs as probabilities of association can be utilized rather than a simple discrete definition of which mRNAs are targets. For example, GSEA analysis would not have been possible simply by defining which mRNAs were targets, it depended on the continuous metric of a rank-ordered list of mRNAs created from LOD scores.

The observation that many of the mRNAs associated with Pum1 belong to a relatively small number of functional groupings is consistent with the RNA operon/regulon model (Keene and Tenenbaum 2002; Hieronymus and Silver 2004; Keene 2007). This model describes how multiple genes with related functions can be

coordinately regulated at the level of the mRNA. Indeed, many aspects of the RNA regulon model are reflected in the results from this study. For example, one key aspect is that post-transcriptional regulation at the level of the mRNA accommodates the multi-functionality of eukaryotic proteins by allowing a single gene to participate in multiple regulons (cis-combinatorial). Thus, although Pum1 targets are enriched for regulators of progression through cell cycle yet we observed no function for Pum1 in cell cycle progression (see below), it is likely that these genes also function in processes that are not related to the normal progression of cell cycle, and these alternate processes may in fact be affected by Pum1 perturbation. It will be interesting to determine whether Pum1 functions in the meiotic cell cycle, especially since PUF proteins are known to function in the germ line of various organisms, including mouse (Lin and Spradling 1997; Zhang, Gallegos et al. 1997; Xu, Chang et al. 2007). The cell proliferation and differentiation related targets of Pum1 may reflect a role for Pum1 in self-renewal of stem cells, the proposed ancestral function of PUF proteins.

Previous studies of RBPs that bind AU-rich elements (AREs) were unable to draw strong conclusions regarding conservation of RNA-protein interactions among species for two main reasons. First, the RBPs in question bind RNA through the widely represented RRM motif, and thus true orthologues between species are difficult to discern (Anantharaman, Koonin et al. 2002). Second, the sequences to which ARE RBPs bind are not generally unique and involve elements that use both sequence and

structure (Lopez de Silanes, Zhan et al. 2004). Neither of these problems present themselves when considering PUF proteins. The PUF HD is extremely well conserved across species, yet each species has relatively few PUF genes. The sequence to which the PUF HD binds is also very well conserved, with a UGUR followed by an AU rich sequence identified in most PUF binding sites, and the eight nucleotide core motif UGUAHAUA being almost identical between Puf3, Pumilio, and Pum1 (Wang, Zamore et al. 2001; Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006). Comparison of target genes of these three proteins revealed that they likely regulate different processes; Puf3 binds messages of genes with mitochondrial function (Garcia-Rodriguez, Gay et al. 2007), but there is no enrichment for genes with mitochondrial functions among targets of Pumilio (Gerber, Luschnig et al. 2006) or Pum1. Targets of Pumilio and Pum1 do not share specific functional relationships, and although there is some conservation of target genes, most of the Pum1 targets and Pumilio targets are different (Figure 11). These observations show that the modules of the PUF HD and cognate binding sequence have remained fixed through evolution, while the identities of target messages have changed phylogenetically. This study provides experimental evidence of a re-wiring process that was predicted through analysis of conservation of potential post-transcriptional regulatory elements, showing how a conserved cis-trans interaction can be evolutionarily re-wired to coordinate the expression of different subsets of genes in different species (Mesarovic, Sreenath et al. 2004; Chan, Elemento et al. 2005; Gerber,

Luschnig et al. 2006; Keene 2007; Halbeisen, Galgano et al. 2008). Many of the genes whose mRNAs are bound by Pum1 are known regulators of gene expression and cellular processes, and thus Pum1 could be described as a regulator of regulators (Mesarovic, Sreenath et al. 2004; Keene 2007; Pullmann, Kim et al. 2007). The GO category containing the most Pum1 targets is "regulation of biological processes," and target genes of Pum1 contain genes that regulate gene expression at the transcriptional, post-transcriptional, and post-translational levels. GO analysis revealed that transcription factor genes and ubiquitin cycle genes are enriched in Pum1 targets, and many of the target genes involved in nucleic acid metabolism encode proteins that bind and process RNA. Pum1 targets are also enriched for GTPase mediated signal transduction genes and other genes involved in signaling pathways that themselves could be described as regulatory as their activation or repression typically results in changes in gene expression. Thus, Pum1 could be described as a regulator of regulators because it associates with genes that regulate multiple levels of gene expression, as well as genes encoding members of signaling pathways that trigger changes in gene expression.

Pum1's function as a regulator of regulators is also evident when observing genes that are not represented in Pum1 targets. The GO categories of "electron transport" and "oxidoreductase" were significantly depleted among Pum1 targets, indicating that Pum1 does not regulate mRNAs involved in processes that are not

regulatory. Pum1 targets did not show significant enrichment of any GO categories involving metabolism (other than nucleic acid metabolism, which is related to transcription and RNA processing), again demonstrating the lack of a role for Pum1 in non-regulatory processes and supporting its role as a regulator of regulators.

After the results from the RIP-Chip analysis of Pum1 were published (Morris, Mukherjee et al. 2008), another group published a very similar study in which they identified mRNAs associated with Pum1 and Pum2 in human HeLa cells (Galgano, Forrer et al. 2008). The results from both studies are largely in agreement despite substantial variations in experimental procedure, showing a high degree of overlap of Pum1 target mRNAs and potential RNA operons, as well as an identical binding motif. The authors also demonstrated that Pum1 targets largely overlap with those of Pum2, indicating that the proteins may have redundant functions (Galgano, Forrer et al. 2008).

4. Mechanisms of Pum1 regulation of associated mRNAs

4.1 Rationale

It has been previously demonstrated that Puf proteins repress expression of target mRNAs, either through repression of translation, enhancement of decay, or a combination of the two. Drosophila Pumilio protein represses translation of Hb mRNA (Wharton, Sonoda et al. 1998), Xenopus Pum2 represses translation of RINGO/SPY mRNA by competing for the translation initiation factor eIF4E (Cao, Padmanabhan et al.; Padmanabhan and Richter 2006), and yeast Puf5 enhances deadenylation of *HO* mRNA through interaction with the CNOT family of deadenylase proteins (Goldstrohm, Hook et al. 2006; Goldstrohm, Seay et al. 2007). Thus, we hypothesized that Pum1 represses target mRNAs through either repression of translation, enhancement of decay, or both. To test these possibilities, we sought to alter the abundance of Pum1 protein using siRNA mediated knockdown and transfection mediated overexpression.

In order to assay poly(A) tail lengths (PATs) of Pum1 target mRNAs, we employed a method referred to as GI tailing. It has been shown that poly-A polymerase can added a limited number of G residues to the end of a poly-A tail (Martin and Keller 1998), and exploiting this enzyme it is possible to add a limited number of G and I residues onto the ends of purified mRNAs. Inosine is added along with guanine in order to prevent potential secondary structure caused by long strings of Gs. A primer specific to the poly-A+GI tails is then used in a reverse transcription reaction, which

unlike priming using oligo d(T), which will anneal at random locations along the PAT, will retain the entire lengths of the PATs by annealing only at the end, where the GI tail has been added. These cDNAs can then be used as a template for PCR using a gene specific forward primer located a few hundred bases upstream of the polyadenylation site and a universal reverse primer specific to the poly-A+GI tails, resulting in a pool of PCR products that represents the distribution of poly-A tail lengths of the mRNA being assayed. A diagram of this method is shown in Figure 12.

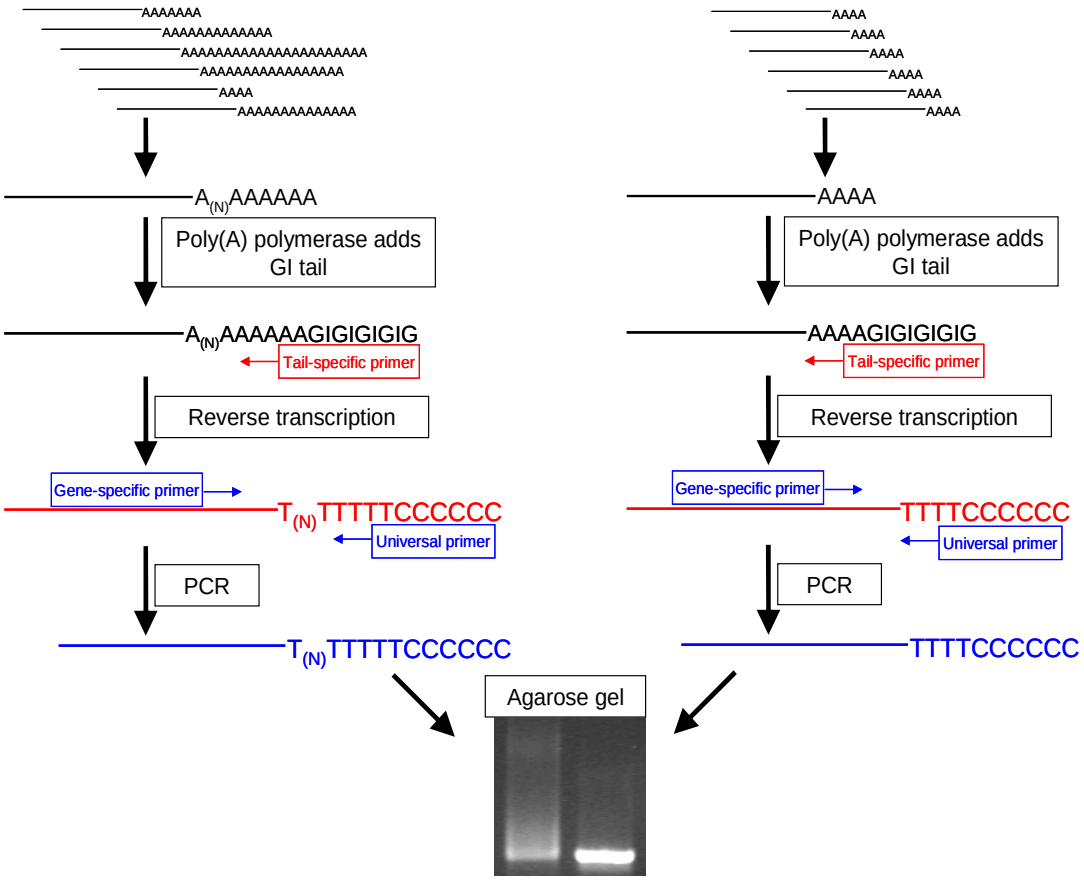


Figure 12: GI tailing PAT assay. Poly(A) polymerase is used to add a limited number of G and I residues to the ends of mRNA poly(A) tails (PATs). A primer specific to these poly(A) + GI tails is used for reverse transcription, then PCR is

performed with a gene specific forward primer and a universal reverse primer. mRNAs with a distribution of short to long PATs will appear as a smear on the gel, while those with a uniform distribution of short PATs will appear as a sharp band.

4.2 Results

4.2.1 siRNA mediated Pum1 knockdown

In order to deplete Pum1 protein, three siRNAs, which we referred to as siRNAs 17, 18, and 19, were obtained and tested in various combinations and various concentrations. Pum1 mRNA levels were assayed in order to determine the optimal conditions for Pum1 knockdown (KD), which was a combination of siRNAs 17 and 19 at a final concentration of 50nM each. These conditions allowed us to consistently achieve a >70% Pum1 mRNA reduction and >90% protein reduction, compared to a control, non-targeting siRNA.

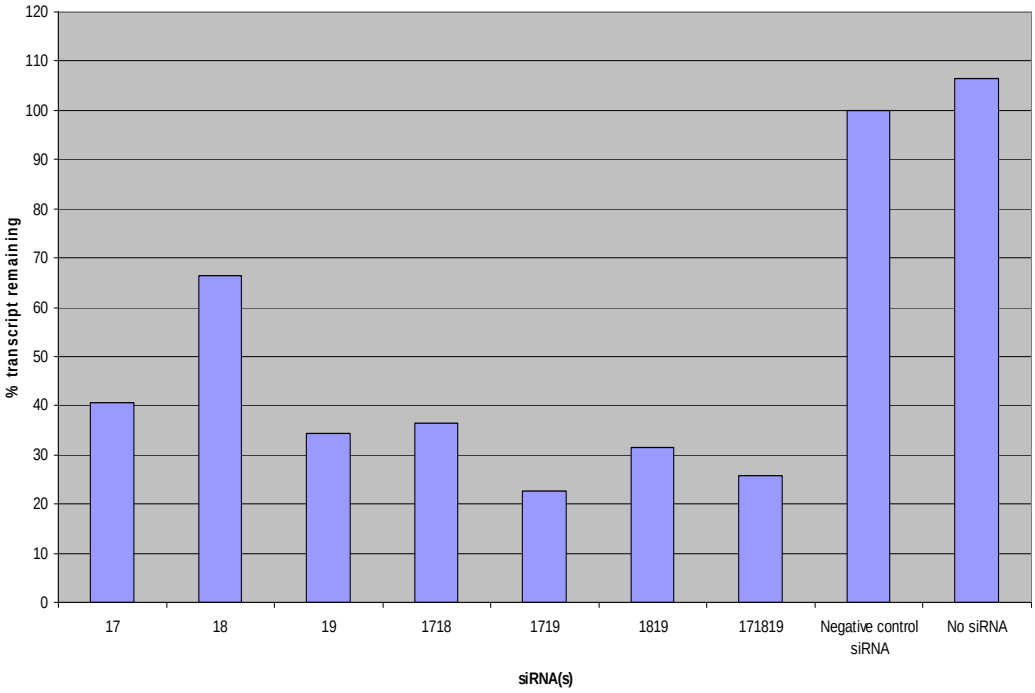


Figure 13: Optimization of Pum1 knockdown. Three siRNAs were tested in all combinations at a final siRNA concentration of 100nM and Pum1 mRNA abundance was assayed by qPCR.

4.2.2 Pum1 overexpression

The ability to overexpress (OE) Pum1 protein was also deemed valuable to study the functions of Pum1, thus a plasmid was created containing the entire coding sequence of the Pum1 protein in the pcDNA3 vector. The ability to overexpress full-length Pum1 protein from this plasmid was confirmed by Western blot. It is noteworthy that this construct contains the entire coding sequence of Pum1, as many other studies of Puf proteins express in vivo or in vitro only the RNA binding PUF domain.

4.2.3 Protein abundance of Pum1 targets after Pum1 perturbation

As it could be argued that protein production is the most relevant consequence of PTR, so we first assayed the protein levels of various Pum1 targets by Western blotting. It was determined that there was no obvious difference in the abundance of PCNA or Cyclin B1 protein in Pum1 KD cells versus control (Figure 14). There also did not appear to be a difference in PCNA protein abundance after Pum1 overexpression (Figure 15).

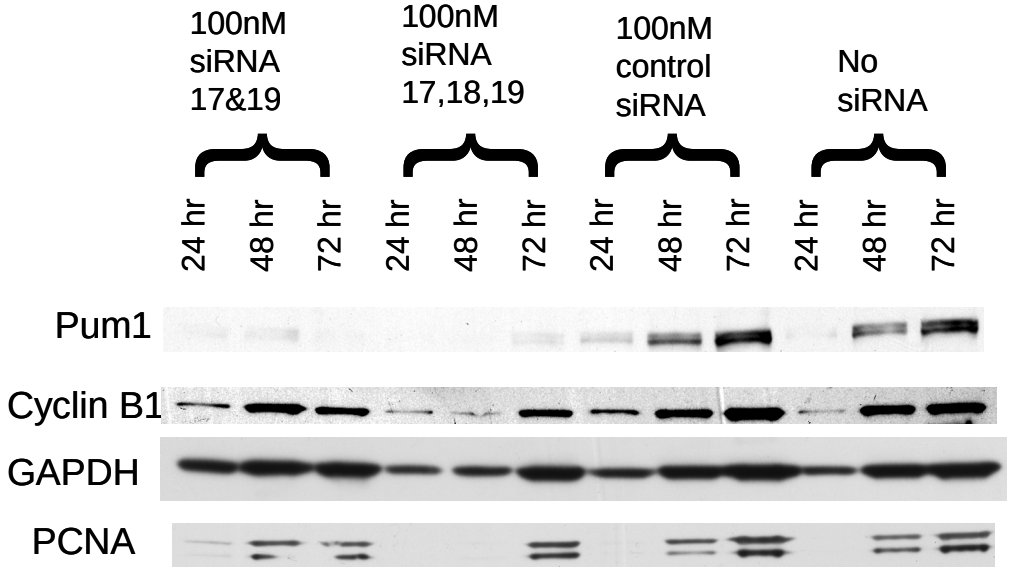


Figure 14: Protein abundance of Pum1 targets during Pum1 knockdown. Pum1 protein was knocked down using the siRNAs and times indicated and protein abundance of the Pum1 targets Cyclin B1 and PCNA and non-target GAPDH were assayed by Western blot.

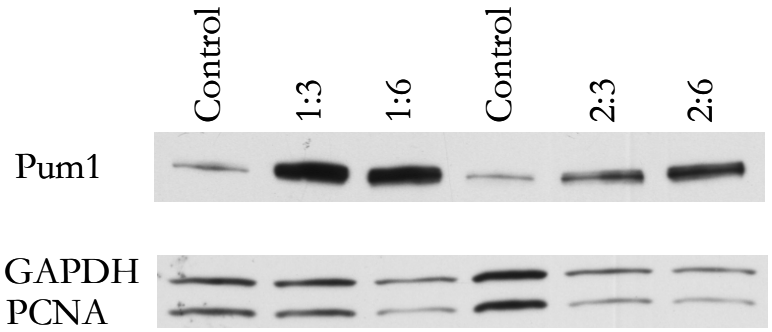


Figure 15: Protein abundance of Pum1 targets during Pum1 overexpression. Pum1 was overexpressed using the indicated ratio of micrograms DNA to microliters of transfection reagent and the protein abundance of the Pum target PCNA and non-target GAPDH assayed by Western blot.

4.2.4 Abundance of mRNA targets after Pum1 perturbation

Pum1 OE affected most mRNA targets in an expected manner, with mRNA abundance decreasing after Pum1 OE (Figure 16). Surprisingly, Cyclin E2 mRNA showed the opposite result, with mRNA abundance increasing. As Pum1 targets a number of transcription factors and RBPs, as discussed previously, this is likely a secondary effect where repression of a repressor led to increased abundance of Cyclin E2 mRNA. As this is also a result from a single replicate it is also possible there was an error, although the qPCR reactions were performed in technical duplicate so this explanation is unlikely.

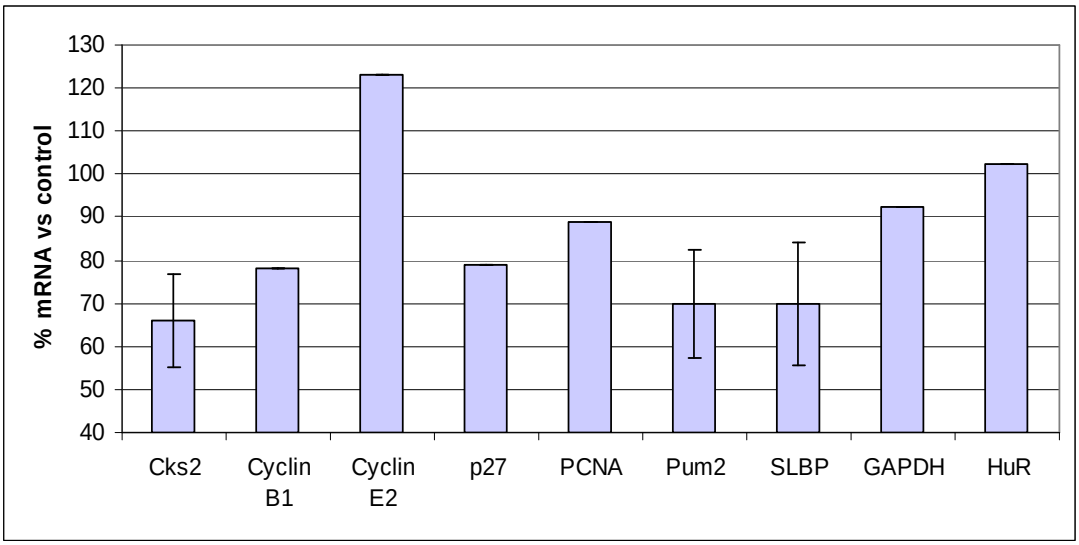


Figure 16: mRNA abundance of Pum1 targets during Pum1 overexpression. Pum1 was overexpressed and the abundance of the indicated target mRNAs and the non-target mRNAs GAPDH and HuR were assayed by RT-qPCR. For samples with error bars, those bars represent the range of biological duplicate experiments.

We then assayed the steady state mRNA levels of Pum1 targets after Pum1 KD. As demonstrated in Figure 17, there was little difference in target mRNA levels that was consistently detectable after Pum1 KD, with the exception of Pum2 and SLBP mRNAs, which were consistently found to be 40-50% more abundant after Pum1 KD. Although a number of other target mRNAs also appeared more abundant after Pum1 KD, the non-target mRNA B2M showed a similar enrichment, indicating that this level of enrichment may be an off target effect.

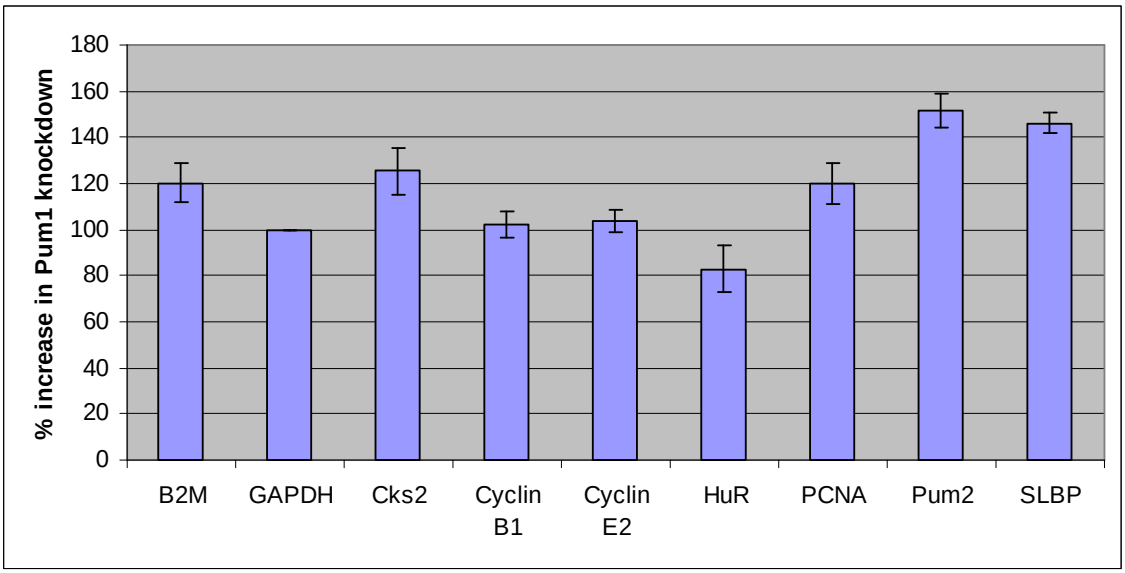


Figure 17: mRNA abundance of Pum1 targets during Pum1 knockdown. Pum1 was knocked down as described in the text and mRNA abundance of targets was assayed by qRT-PCR. B2M, GAPDH, and HuR are non-targets, all others are targets. Bars represent SEM of three replicates.

4.2.5 Role of Pum1 in decay of target mRNAs

We next sought to determine whether Pum1 enhances decay of associated mRNAs, as has been shown for other PUF proteins (Olivas and Parker 2000; Goldstrohm, Hook et al. 2006). Thus, we knocked down Pum1 protein using siRNA, then assayed decay rates of target messages. Extent of Pum1 protein depletion was determined to be approximately 70-95% by Western blot for all decay experiments (Fig 18). To determine mRNA decay rates, Actinomycin D was used to inhibit transcription and qPCR was performed to determine the percent of transcripts remaining at multiple time points after treatment, as normalized to GAPDH and averaged across three biological replicates. An exponential decay curve was fit to the mean of these

measurements and half-lives of messages were determined based on the equation describing this curve (Figure 19). Control experiments were performed with an siRNA not known to target any mRNAs.

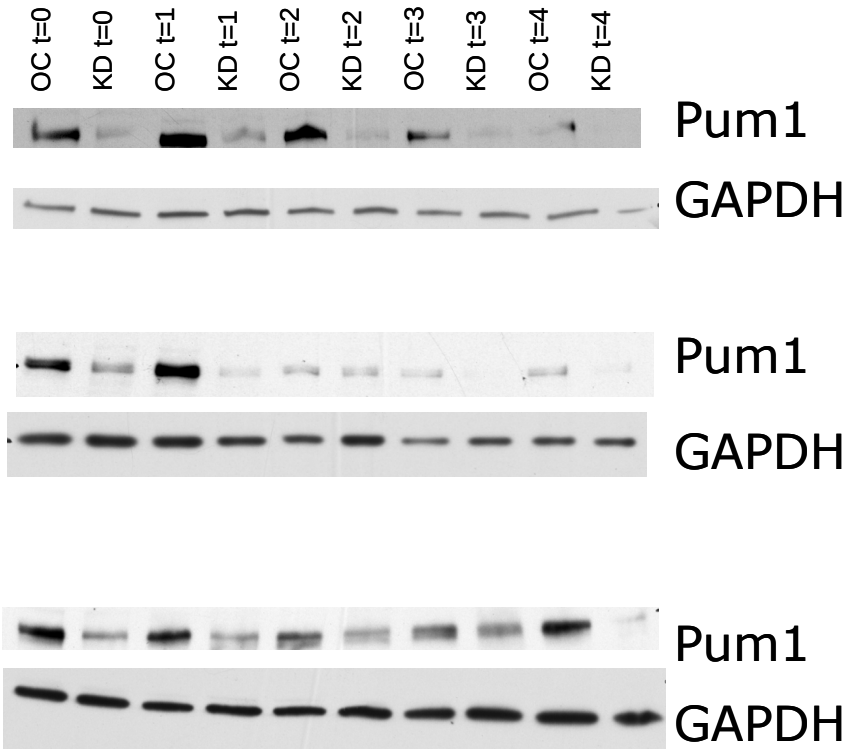


Figure 18: Pum1 protein knockdown during mRNA decay assay. Abundance of Pum1 was assayed by Western blot in samples used for mRNA decay rate determinations in next Figure. OC=Oligo control, KD=Pum1 knockdown, t=hours after addition of Actinomycin D.

All but one of the Pum1 target mRNAs assayed showed increased stability during Pum1 knockdown, although there was a range in the degree of increased stability. As Pum1 protein has been shown to interact with CNOT8 protein (20) (Goldstrohm, Hook et al. 2006), a member of the CCR4-NOT deadenylase complex, it

will be interesting to determine in future studies whether Pum1's effect on stability of associated mRNAs is at least partially mediated by deadenylation via this complex.

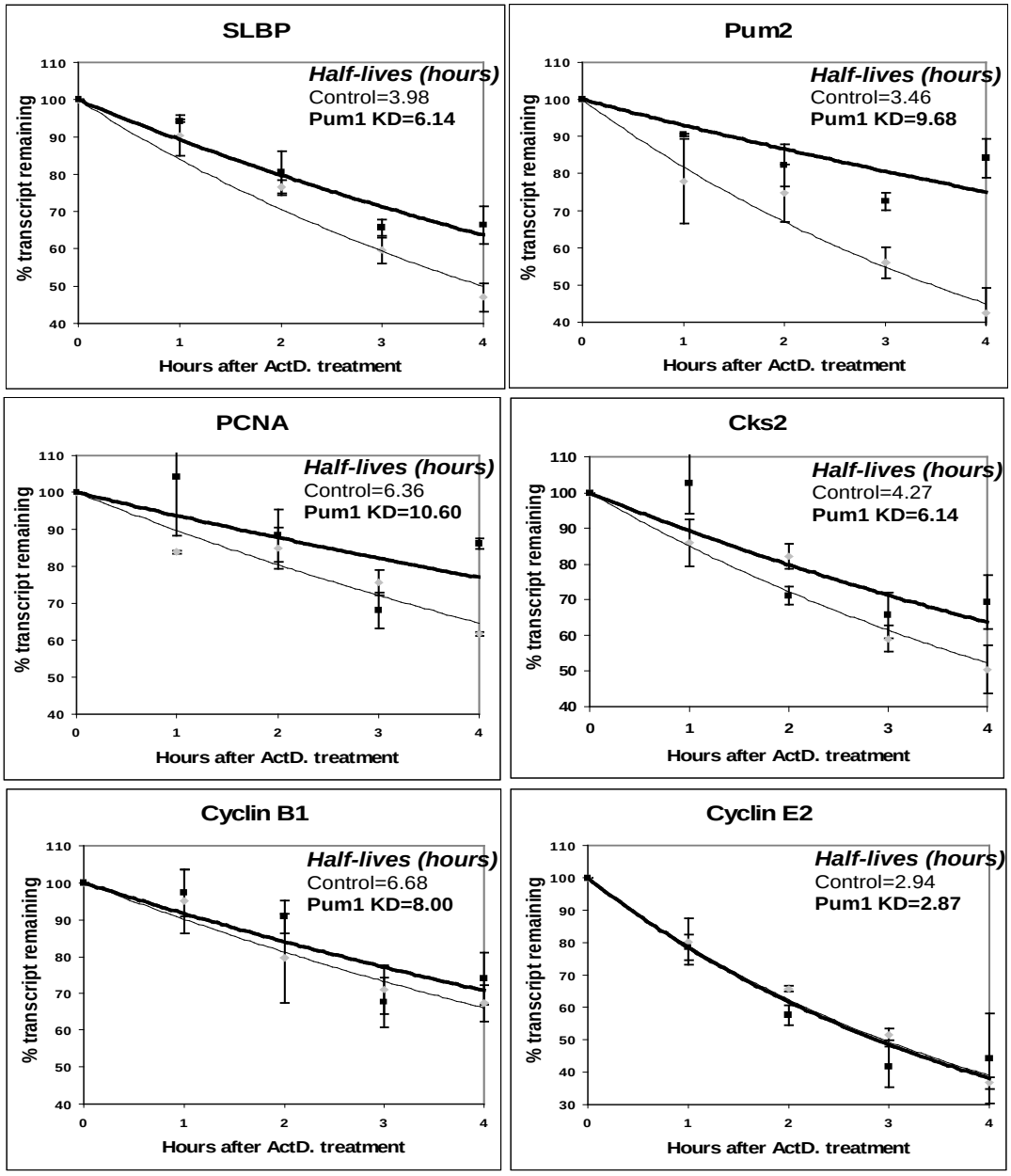


Figure 19: Decay rates of Pum1 target mRNAs during Pum1 knockdown. Decay rates of Pum1 target mRNAs, normalized to GAPDH, as determined by RT-qPCR after Pum1 knockdown followed by treatment with Actinomycin D to inhibit transcription. Black boxes (bold curve) represent Pum1 knockdown, grey diamonds (normal curve) represent control. Error bars represent SEM of three biological replicates. X-axis represents hours after addition of Actinomycin D, y-axis represents percent transcript remaining.

A pilot experiment was performed to determine if Pum1 overexpression would result in more rapid decay for target mRNAs, as would be predicted by the knockdown experiment. This experiment yielded the opposite result (Fig 20), with target mRNAs decaying less rapidly. This result was likely a dominant negative artifact of Pum1 overexpression, possibly caused by saturating the binding partners of Pum1 with Pum1 molecules not bound to mRNA.

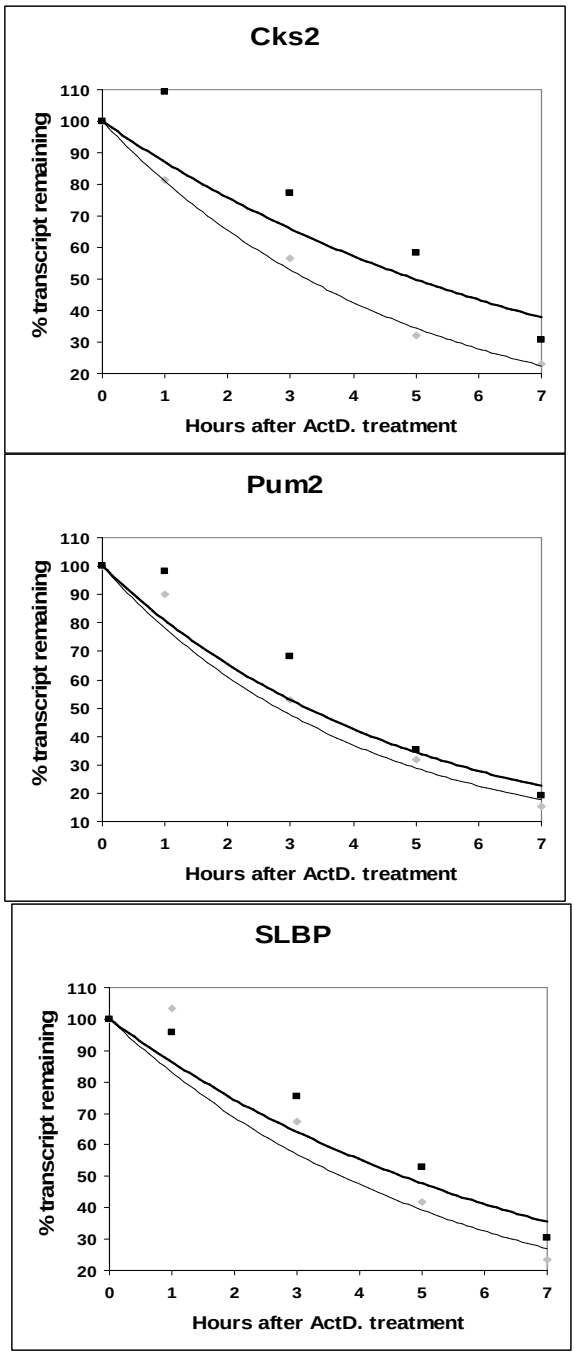


Figure 20: Pum1 target mRNA decay during Pum1 overexpression. Decay rates of Pum1 target mRNAs, normalized to GAPDH, as determined by RT-qPCR after Pum1 overexpression followed by treatment with Actinomycin D to inhibit transcription. Black boxes and bold lines are Pum1 overexpression, gray diamonds

and normal lines are control. X-axis represents hours after addition of Actinomycin D, y-axis represents percent transcript remaining.

4.2.6 Role of Pum1 in poly(A)-tail length of target mRNAs

In order to determine whether Pum1 enhances decay of associated mRNAs via enhancement of deadenylation we assayed poly-A tail (PAT) lengths of Pum1 target mRNAs after knockdown of Pum1 protein and in Pum1 IPs using the GI tailing method described in Rationale. Figure 21 is an explanation of how data from this assay are analyzed and presented.

In order to confirm that PCR products indeed represented PATs, we treated RNA samples with RNAse H either plus or minus oligo d(T), as seen in Figure 22. As RNAse H degrades RNA/DNA duplexes, all PAT tails in samples treated with oligo d(T) and RNAseH should display very short PATs, basically representing only the A residues that were not hybridized to a T due to the position at which the oligo d(T) hybridized. Indeed, in all samples treated with RNAse H and oligo d(T) smears representing PATs collapsed into sharp bands, confirming that PCR products represented bona fide PATs.

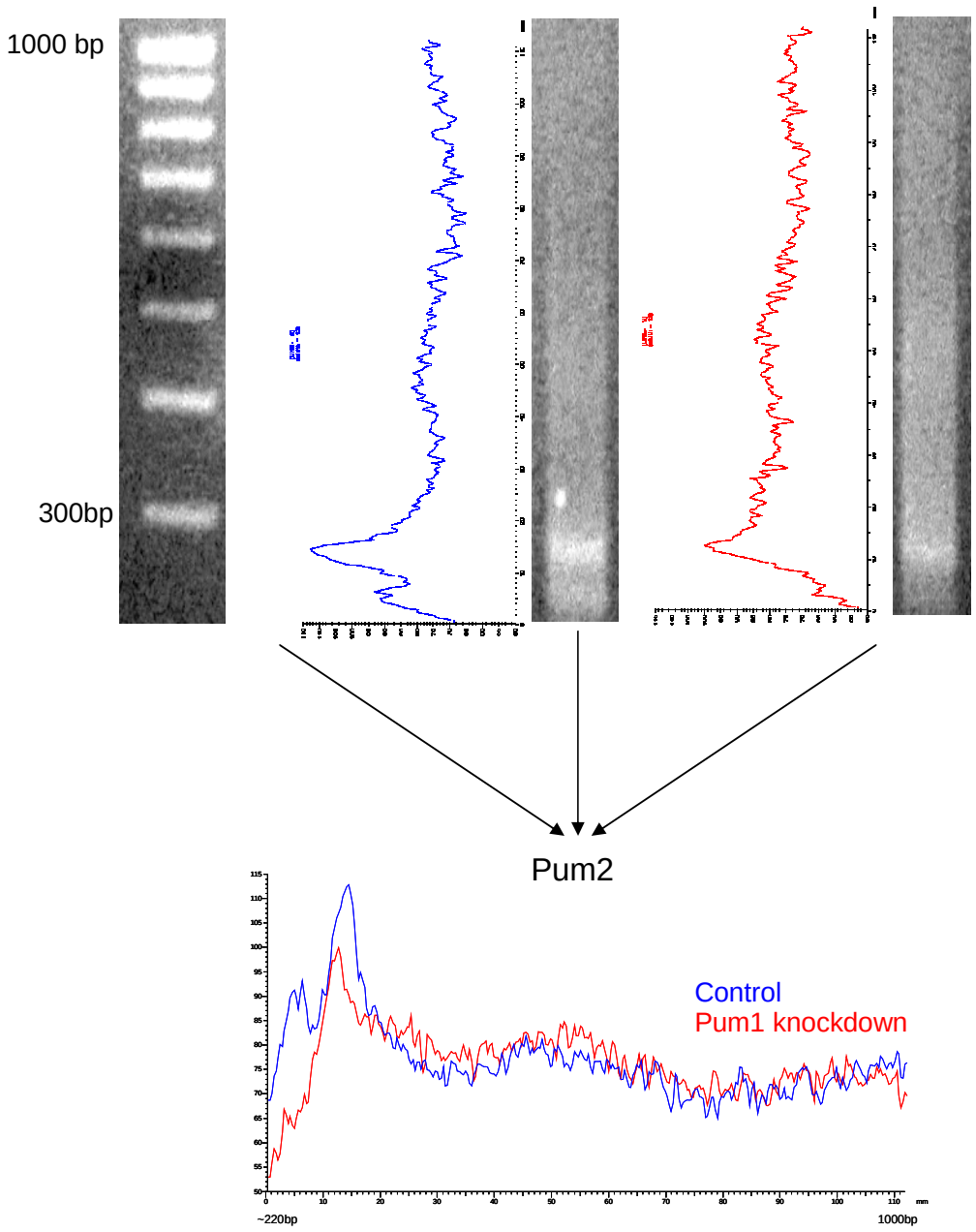


Figure 21: Explanation of GI tailing PAT assay data presentation. Signal intensity is converted to graph form by ImageQuant, and comparison of two conditions allows determination of relative PAT lengths. Signal intensity is represented on the y-axis, while approximate position in the agarose gel is indicated on the x-axis.

If Pum1 normally enhances deadenylation of associated mRNAs, it would be expected that knockdown of Pum1 would result in Pum1 target mRNAs having longer PATs. Indeed, this is what was seen when we assayed PAT length of various Pum1 target mRNAs after Pum1 knockdown, as shown in Figure 23. Pum1 target mRNA PATs were assayed after 48 hours of Pum1 knockdown and 4 hours of treatment with Actinomycin D. PCR products were resolved on 2.5% agarose gels, and intensities of PAT signals determined using ImageQuant software. PATs of Pum2, Cks2 and Cyclin B1 mRNA were relatively longer after Pum1 knockdown, while those of PCNA (not shown) and the control mRNA GAPDH were not affected.

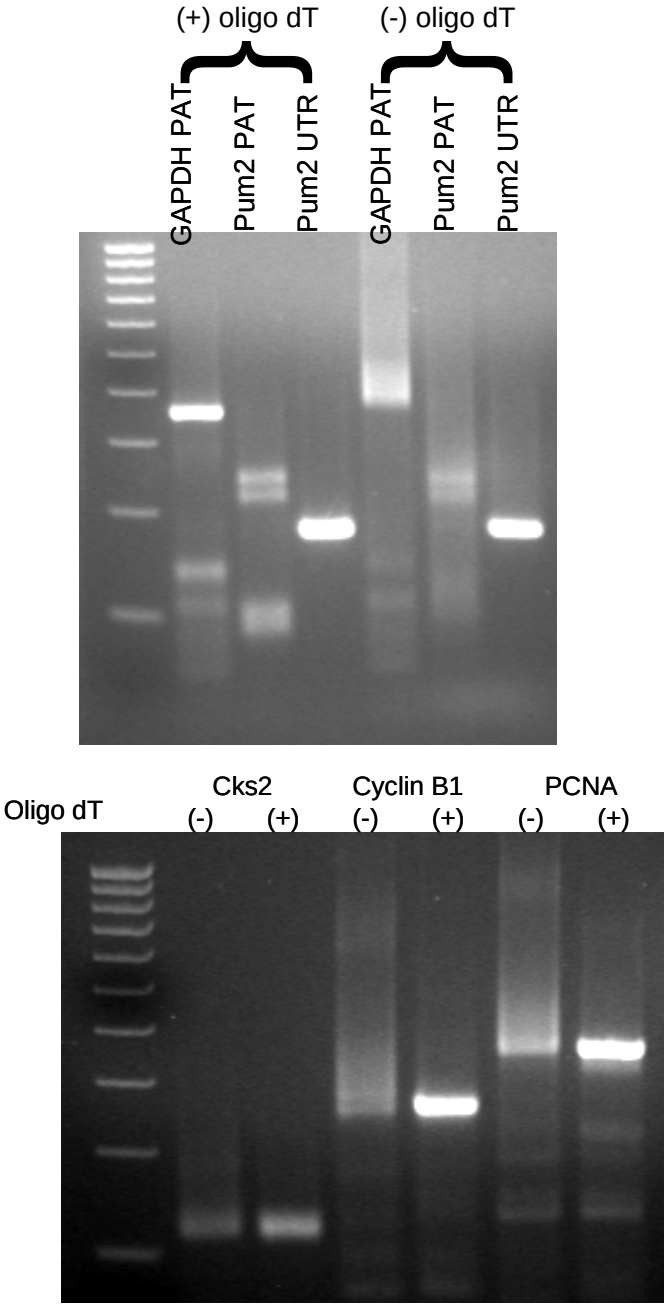


Figure 22: RNAse H control of PATs. RNA samples were treated with RNAse H plus and minus oligo dT in order to confirm that gel smears seen in indicated PCR products are PATs. Pum2 UTR is a control PCR that is not a PAT.

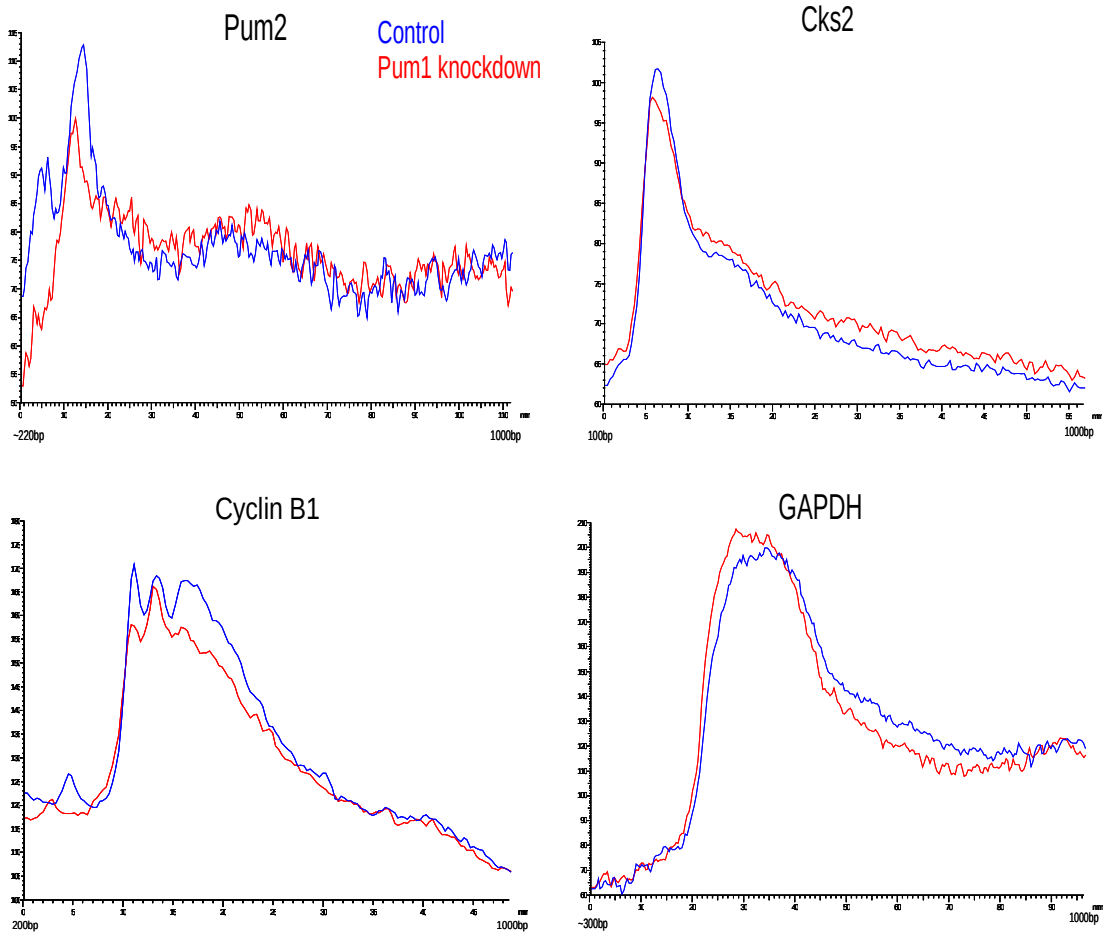


Figure 23: PATs of Pum1 targets during Pum1 knockdown. Pum1 was knocked down and PAT length of mRNA targets and the control GAPDH were assayed via GI tailing, demonstrating that PAT lengths of Pum1 target mRNAs are relatively longer during Pum1 knockdown. Signal intensity is represented on the y-axis, while approximate position in the agarose gel is indicated on the x-axis.

If Pum1 normally enhances deadenylation of associated mRNAs, it would also be expected that mRNAs found associated with Pum1 will have shorter PATs than the total population of that same species of mRNA. Although this might not be immediately apparent, given the large number of Pum1 target mRNAs it is unlikely that

every molecule of every Pum1 target mRNA will be bound by a molecule of Pum1 protein, thus those that are bound by Pum1 represent a subpopulation that should have shorter PATs than the total population. We found that Pum1 target mRNAs did indeed have shorter tails than the same mRNA in total RNA for all Pum1 targets assayed except Cyclin B1. A caveat to this result is that we were unable to assay a control RNA due to the very low abundance of non-target mRNAs in the Pum1 IP.

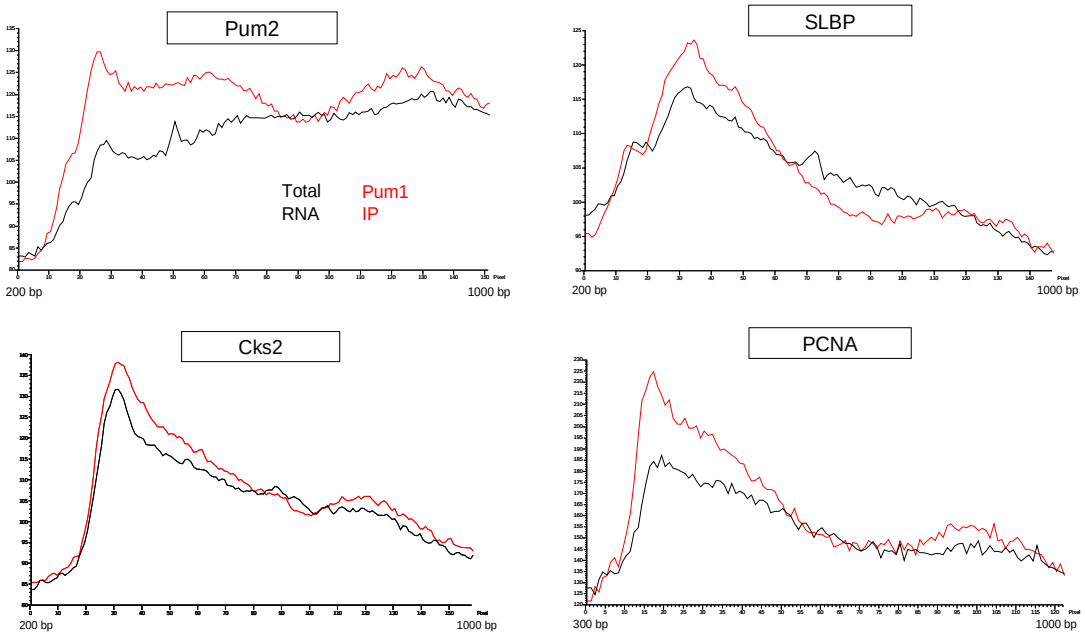


Figure 24: PAT length of Pum1 targets in Pum1 IP versus total RNA. Pum1 associated and total RNA were isolated and PAT length of Pum1 mRNA targets were assayed via GI tailing, demonstrating that PAT lengths of Pum1 target mRNAs are relatively shorter in the Pum1 IP. Signal intensity is represented on the y-axis, while approximate position in the agarose gel is indicated on the x-axis.

4.2.7 Role of Pum1 in translation of associated mRNAs

Puf proteins have also been hypothesized to be involved in repression of translation, thus we sought to test if Pum1 was involved in repression of translation of

target mRNAs by determining if the translational status of these mRNA changed during Pum1 perturbation. We used polysome profiling via sucrose gradients in order to assay translational status. Briefly, cell lysates were prepared in the presence of cycloheximide and spun in 15-50% sucrose gradients. These gradients were then collected in 10 fractions, with the A254 profile being constantly recorded while the fractions were collected. By determining the relative abundance of an mRNA in these fractions one can speculate its translational status, especially by comparing two conditions. Comparing Pum1 overexpressing cells to control cells showed no change in translational status of the Pum1 target mRNAs shown in Figure 25, particularly none greater than that seen for the control mRNA β2M. Due to technical difficulties in scaling up Pum1 siRNA mediated knockdown we were unable to assay translational status of Pum1 targets after Pum1 knockdown. However, changes in poly(A) tail length and decay rates of mRNA is typically accompanied by changes in translational status, thus while Pum1 knockdown is likely to result in altered translational status it is often a "chicken or the egg" argument trying to determine whether enhancement of decay leads to repression of translation or vice versa. Further experiments will be needed to determine whether Pum1 has a role in repression of translation, and whether this role is correlative or distinct from its role in deadenylation and decay of target mRNAs.

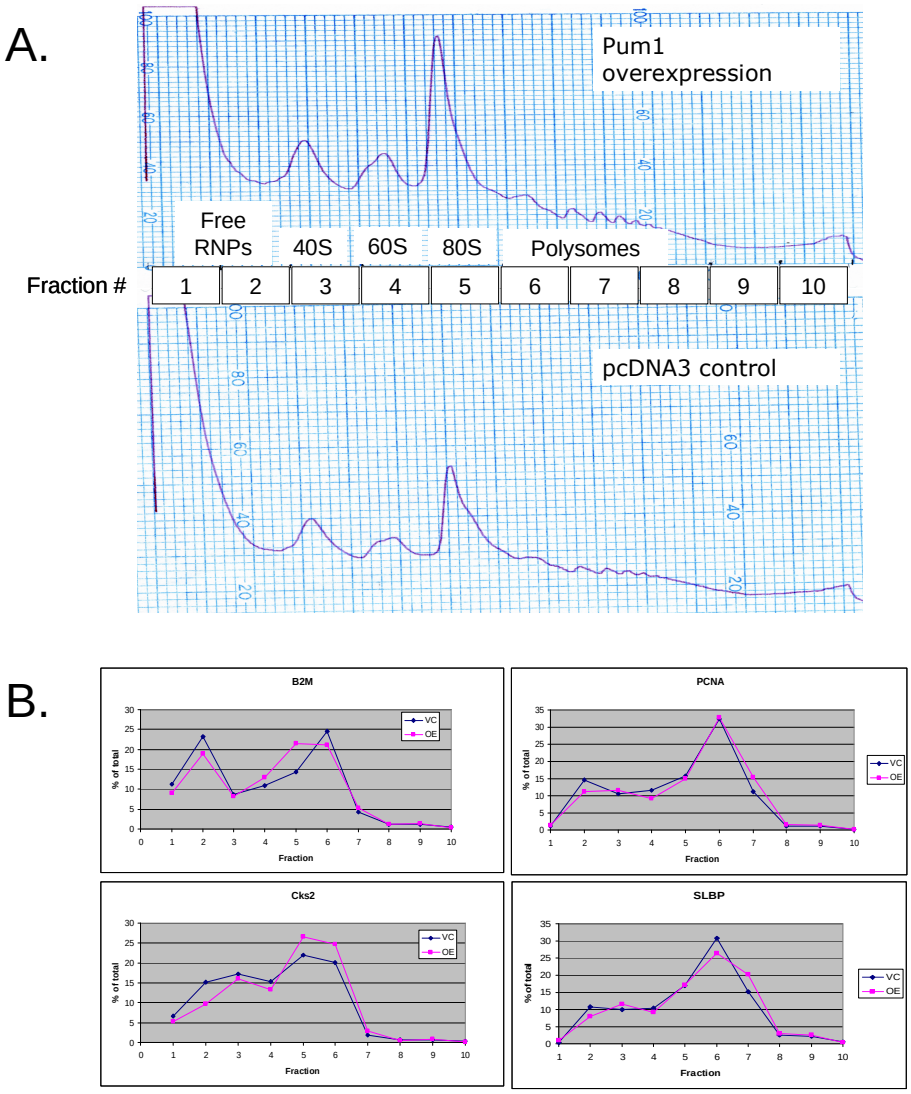


Figure 25: Translation profiles of Pum1 target mRNAs during Pum1 overexpression. Gradient A254 profiles for Pum1 overexpression and control cells are shown in (A), and the relative abundance of the indicated Pum1 target mRNAs and non-target mRNA B2M as determined by RT-qPCR in (B). Fraction numbers from (A) are shown on the x-axis in (B), while percent of total signal represented by each fraction is shown on the y-axis.

4.3 Methods

4.3.1 Pum1 knockdown

Pum1 knockdown was performed using a mixture of 2 Ambion Silencer siRNAs targeting the Pum1 mRNA and siPort NeoFx reagent (Ambion) for transfection, following manufacturer's instructions. Protein knockdown of approximately 70-95% was confirmed by Western blotting for all experiments. Assays were performed 40-48 hours after transfection of siRNAs. A non-targeting siRNA (Ambion) was used as a negative control.

4.3.2 Pum1 overexpression

A plasmid overexpressing the entire Pum1 protein was constructed by cutting the Pum1 open reading frame from IMAGE clone 3528160 using the restriction enzymes EcoRI and XhoI and ligated into pcDNA3 vector which was cut by the same enzymes. Transfection was performed using Expressfect transfection reagent from Denville Scientific according to manufacturer's instructions.

4.3.3 mRNA decay rate determination

mRNA abundance of targets messages was determined by RT-qPCR at hourly time points after addition of 5 ug/ml Actinomycin D. The non-target mRNA GAPDH was used for normalization. Exponential decay curves were fit to points representing the mean of three biological replicates, and half-lives were calculated based on the equations describing these best fit curves.

4.3.4 Poly(A) tail (PAT) assays

PAT lengths were determined using the Poly(A) Tail Length Assay Kit from USB corporation, according to manufacturer's instructions. Products were resolved on 2.5% agarose gels and stained with ethidium bromide. Quantification was performed using ImageQuant software.

4.3.5 Polysome gradients

Polysome gradient assays were performed essentially as described previously (Bradrick, Dobrikova et al. 2007).

For RT-qPCR, RNA was extracted from 100 microliters of each fraction and resuspended in 10 microliters of water, and then 5 microliters of this RNA was used for each RT reaction.

4.4 Discussion

While mechanisms of repression of target mRNAs by PUF proteins have been determined in a variety of species (Cao, Padmanabhan et al.; Wharton, Sonoda et al. 1998; Goldstrohm, Seay et al. 2007), this study represents the first mechanistic study of a mammalian PUF protein. Although it was unexpected to discover that Pum1 does not appear to alter protein abundance of its associated mRNAs there are a number of potential explanations for this result. Pum2, although expressed at relatively low levels, may be compensating for loss of Pum1 by binding to and repressing translation of Pum1 target mRNAs, as is likely given the high degree of overlap of Pum1 and Pum2 target

mRNAs and Pum1 and Pum2 consensus binding sites (Galgano, Forrer et al. 2008). Also, given the large number of mRNAs encoding both positive and negative regulators of gene expression that are associated with Pum1, it is possible that the gene expression network is able to correct itself via de-repression of these genes. It is also possible that changes in protein abundance were too small to detect, especially given that Western blotting is sub-optimal for detecting small changes in protein abundance. This explanation is also supported by the mRNA abundance determinations after Pum1 KD, where differences seen were of a small magnitude and likely only detectable because we assayed with the very sensitive method of qPCR.

We were able to confirm the previous hypothesis that Pum1 enhances decay of associated mRNAs (Goldstrohm, Hook et al. 2006), and in addition demonstrated that this is likely achieved through Pum1's enhancement of deadenylation. Further experiments will be needed to determine what proteins are associated with Pum1 that may catalyze deadenylation and which of these proteins are necessary for Pum1's effect on deadenylation and decay. It was previously shown in vitro that a portion of the human Pum1 protein can bind a human deadenylase, CNOT8 (Goldstrohm, Hook et al. 2006), and although we have been unable to confirm this result in vivo it is likely that Pum1 interacts with at least one member of the relatively large family of CNOT deadenylase proteins.

A significant but somewhat unintuitive result seen in this study is that mRNAs associated with Pum1 RNPs have relatively shorter poly-A tails than the total population of the same species of mRNA. This result supports the hypothesis that not every copy of a single species of mRNA has the same "life," meaning that it is not necessarily bound by the same set of trans factors and thus may undergo differential localization, translation, and decay (Moore 2005; Morris, Mukherjee et al. 2009). When an assay is performed to analyze a single species of mRNA, it is in actuality an average across all copies of that species of mRNA that is being assayed, and thus procedures such as RIP-Chip may be useful for more detailed mechanistic assays by allowing isolation of a sub-population of mRNA whose behavior should be more uniform.

The GI-tailing PAT assay used in this study is a somewhat new technique for determining poly-A tail length. While it is tempting to estimate absolute PAT length based upon results from this assay, there are a large number of assumptions made that must be controlled for in order to make these types of judgments. For example, locations of polyadenylation signals and the exact start sites of the poly-A tail in the 3'UTR of mRNAs are often somewhat ambiguous, and often times multiple polyadenylation sites are used. For example, in this study it appears that there are two polyadenylation sites within close proximity in the Pum2 3' UTR, resulting in two bands of mostly deadenylated product. Sequencing data (not shown) supported the conclusion that these two products indeed represent two distinct polyadenylation sites.

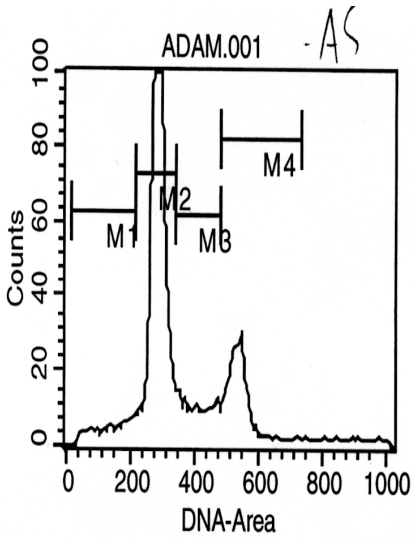
Due to the various caveats involved in the GI tailing PAT assay we avoided estimating actual lengths of poly-A tails, but instead made relative comparisons of the lengths of the poly-A tails of the same species of mRNA under different conditions.

5. Pum1's role in cell cycle

5.1 Rationale

PUF proteins have been hypothesized to have an ancestral role in self-renewal of stem cells, a hypothesis proposed after PUF proteins were found to affect stem cell self renewal and cell cycle in numerous species including C. elegans (Crittenden, Bernstein et al. 2002), Drosophila (Lin and Spradling 1997; Forbes and Lehmann 1998), Dictyostelium (Souza, da Silva et al. 1999), planaria (Salvetti, Rossi et al. 2005) and Xenopus (Padmanabhan and Richter 2006).

Results from Pum1 RIP-Chip experiments in HeLa and Jurkat cells also indicated a role for Pum1 in cell cycle. "Progression of cell cycle" was an enriched GO category of both analyses, and that result in combination with PUF proteins' proposed ancestral role in self-renewal of stem cells led us to hypothesize that Pum1 has a role in cell cycle in HeLa cells.



File: ADAM.001 Sample ID:
Acquisition Date: 09-Dec-05 Gate: G1
Gated Events: 9499

	Marker	Left, Right	Events	% Gated	Mean	CV
	All	0, 1023	9499	100.00	344.68	38.62
	M1	15, 216	603	6.35	141.24	37.55
G1	M2	216, 340	5872	61.82	283.23	7.41
S	M3	340, 482	1225	12.90	409.17	10.48
G2	M4	480, 734	1684	17.73	536.69	7.97

Figure 26: Explanation of FACS based cell cycle analysis. Cells are gated into four groups based on DNA content, and which group corresponds to which cell cycle stage is shown. X-axis represents DNA content, y-axis represents number of cell. AS=Asynchronous, meaning this is the profile of normally growing HeLa cells.

5.2 Results

5.2.1Abundance of Pum1 protein and target mRNAs during cell cycle

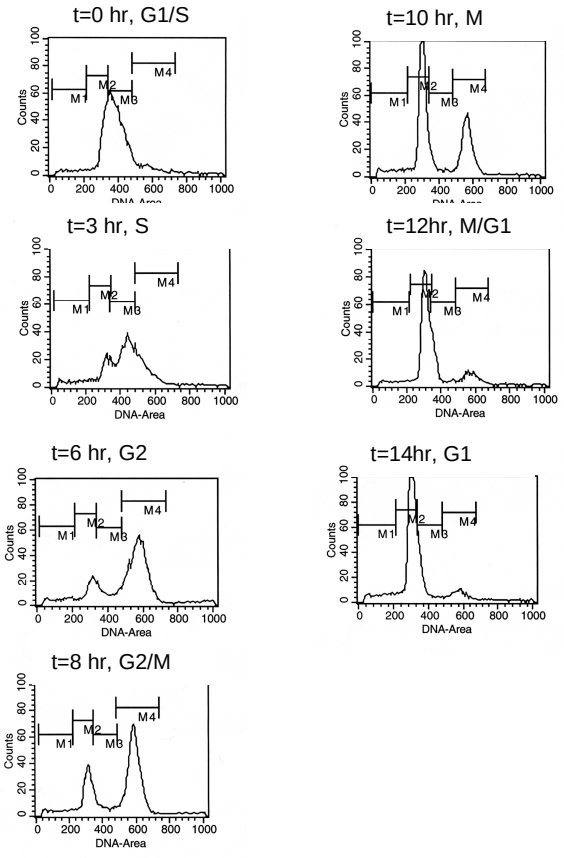


Figure 27: Cell cycle synchronization. FACs profiles of cells synchronized at the G1/S boundary by double thymidine block. Times indicate hours after release from block, letters indicate stage of the cell cycle.

As results from Pum1 target mRNA analysis indicated a role for Pum1 in cell cycle, we sought to determine whether Pum1 has a role in cell cycle in HeLa cells. HeLa cells were synchronized using a double thymidine block, which causes cells to halt cell cycle at the G1/S boundary and then continue through about 2 synchronous cell cycles once released from the block (Whitfield, Sherlock et al. 2002). Cell cycle stage was

determined by FACS as shown in Figure 26. Synchronization of cells was determined to be ~85% by FACS analysis, and Western markers of cell cycle also demonstrated highly synchronous cell cycle: rapid and abundant de-phosphorylation of Cdk1 and decay of Cyclin A during mitosis (Figure 27). No obvious changes in Pum1 protein abundance were observed during cell cycle in three independent biological replicates, indicating that change in Pum1 protein expression and thus repression of target mRNAs is not occurring during cell cycle in HeLa cells.

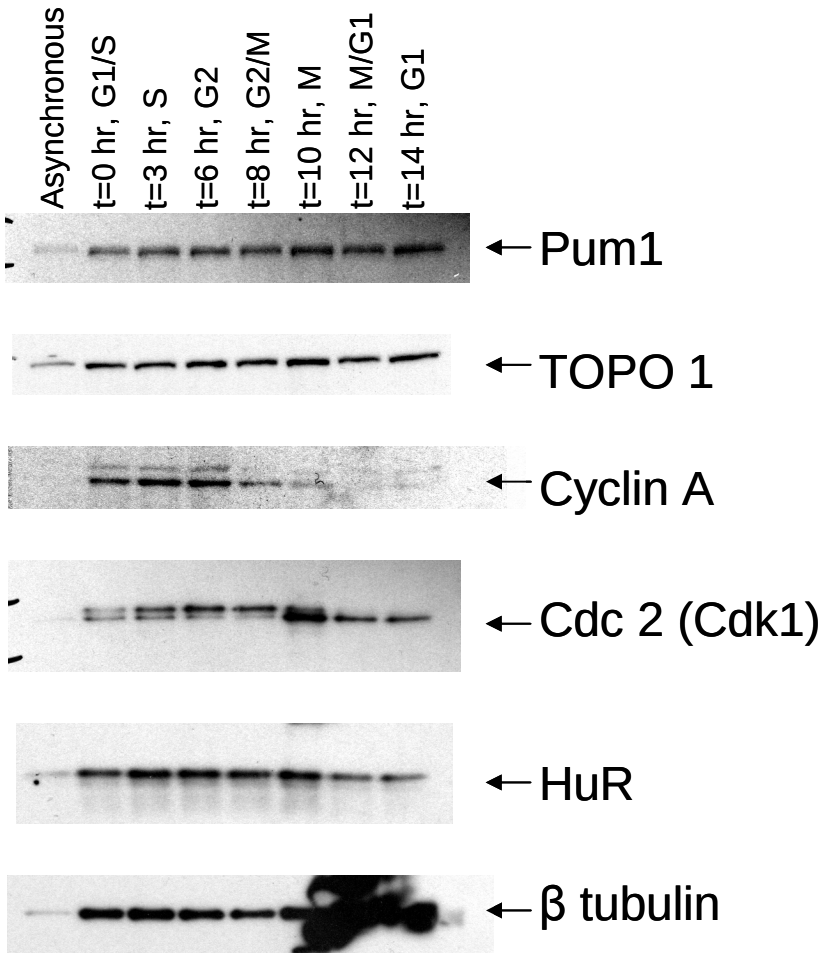


Figure 28: Protein abundance of Pum1 during the cell cycle. Western blot of Pum1 and indicated proteins during the cell cycle is shown. Cyclin A and Cdc2 are markers of cell cycle stage, TOPO1 and β-tubulin are loading controls. Hours after release from G1/S bock and approximate cell cycle stage are indicated.

5.2.2 Effect of Pum1 knockdown on cell cycle progression

While expression of Pum1 protein did not appear to be cell cycle regulated, it was still possible that Pum1 affects cell cycle in a manner independent of its protein abundance. Thus, we sought to determine if cell cycle progression was altered in cells

where Pum1 protein had been depleted via siRNA. Cell cycle profiles of asynchronous Pum1 knockdown cells revealed identical cell cycle profiles to cells treated with a control siRNA or no siRNA despite a large degree of protein depletion (Figure 29), providing compelling evidence that Pum1 does not have a role in normal progression of cell cycle in HeLa cells.

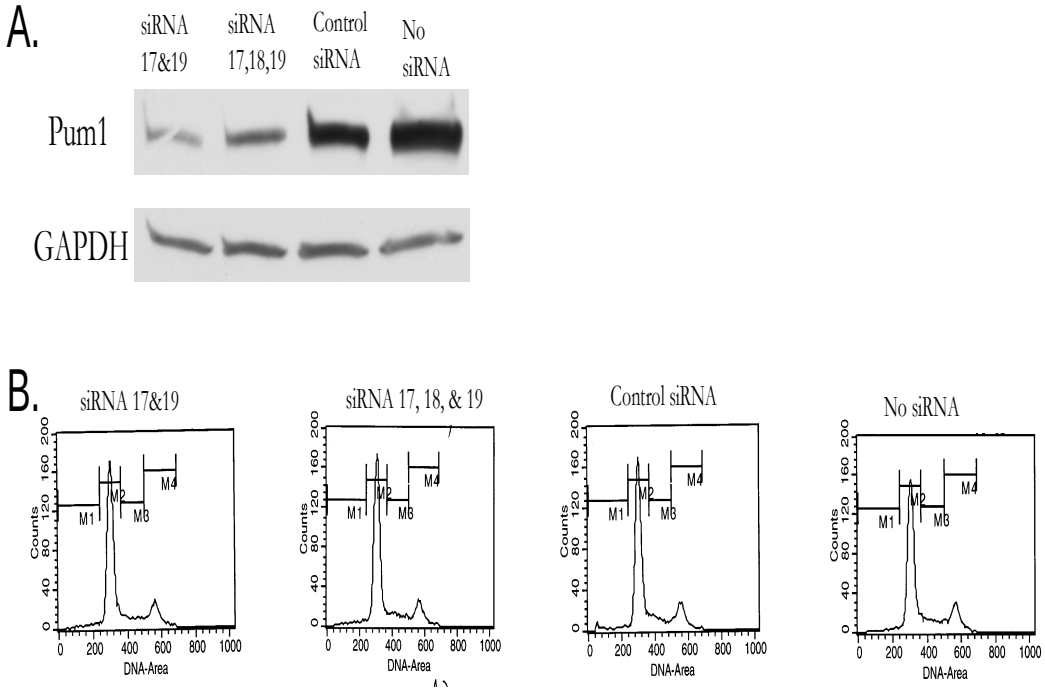


Figure 29: Cell cycle progression during Pum1 knockdown. Western blot in (A) demonstrates extent of Pum1 protein depletion, FACs profiles in (B) indicates no difference in cell cycle profile between Pum1 knockdown and control cells.

5.3 Methods

5.3.1 Cell cycle synchronization

Cell cycle synchronization was performed using a double thymidine block as described previously (Whitfield, Sherlock et al. 2002).

5.3.2 Cell cycle phase determination

Cells were fixed in 70% ethanol and DNA stained with propidium iodide. FACs analysis was used to determine DNA content and thus cell cycle phase.

5.4 Discussion

The result that Pum1 does not seem to alter progression of normal cell cycle in HeLa cells was not entirely unexpected. As the proposed ancestral function of PUF proteins is in self-renewal of stem cells, it is possible that Pum1 only affects cell cycle in those cells and not in HeLa cells. It is also possible that Pum1 acts independently of cell cycle to promote this self-renewal, and the enrichment seen for cell cycle related categories may indicate a role for many cell cycle genes in self-renewal of stem cells. It should also be noted that while HeLa cells have been advantageous for studying concepts of mammalian cell cycle, their vast difference from any normal cell in the human body is part of what makes them useful for scientists, and thus Pum1 may be important for cell cycle in vivo but not necessarily in HeLa cells.

Numerous studies have demonstrated a role for PUF proteins in architecture and function of neuronal cell types (Schweers, Walters et al. 2002; Mee, Pym et al. 2004; Menon, Sanyal et al. 2004; Vessey, Vaccani et al. 2006; Muraro, Weston et al. 2008; Fiore, Khudayberdiev et al. 2009; Menon, Andrews et al. 2009), thus is also possible that Pum1 represses cell cycle related and other mRNAs to achieve these functions. For example, rat Pum2 is involved in dendrite morphogenesis and synapse function (Vessey,

Schoderboeck et al.), and Drosophila Pumilio regulates neuronal excitability and dendrite morphogenesis (Schweers, Walters et al. 2002; Mee, Pym et al. 2004; Menon, Sanyal et al. 2004; Muraro, Weston et al. 2008; Fiore, Khudayberdiev et al. 2009; Menon, Andrews et al. 2009). Future studies will be needed on different types of cells, including stem cells and neurons, to determine biological functions of Pum1 and how they relate to its mRNA targets.

6. Subcellular localization and dynamics of Pum1

6.1 Rationale

In addition to identification of mRNA targets and effects on these targets, knowledge of the subcellular localization and dynamics of an RBP can be useful for determining the functions of that RBP. Many RBPs have a distinct subcellular localization that often changes upon changing cellular conditions, such as addition of a stressor (Anderson and Kedersha 2006; Anderson and Kedersha 2008). Perhaps the most obvious demonstration of this effect is the formation of stress granules and increase in size and number of processing bodies (p-bodies) upon addition of stress to many cell types (Anderson and Kedersha 2006; Anderson and Kedersha 2008). Stress granules are large cytoplasmic aggregates containing numerous RBPs, mRNA, the 40S ribosomal subunit and a number of initiation factors (Anderson and Kedersha 2006; Anderson and Kedersha 2008). Stress granules are found rarely in normally growing cells but induced rapidly after addition of many types of stress, such as oxidative stress induced by arsenite, heat shock, and UV irradiation (Anderson and Kedersha 2006; Anderson and Kedersha 2008). P-bodies also contain many RBPs and mRNA, however instead of the 40S subunit and intiation factors they contain a number of mRNA decay enzymes, such as Dcp1a and Ddx6 (Anderson and Kedersha 2006; Anderson and Kedersha 2008). P-bodies are typically found in normally growing cells, however they become larger and more numerous upon addition of stress, and can be observed to

physically interact with stress granules (Anderson and Kedersha 2006; Anderson and Kedersha 2008). It has been proposed that mRNA may be passed between stress granules and p-bodies via RBPs, with stress granules being storage sites for translationaly inactive mRNAs and p-bodies being sites of mRNA decay (Anderson and Kedersha 2006; Anderson and Kedersha 2008). Human Pum2 protein was found to localize to stress granules and in fact be necessary for their formation (Vessey, Vaccani et al. 2006), thus we sought to determine if Pum1 behaved similarly.

Previous work from the Keene lab has demonstrated that the targets of the RBP HuR change during the dynamic condition of T-cell activation (Mukherjee, Lager et al. 2009). We demonstrate below that Pum1 changes localization during cell stress, and as this is a dynamic condition that is known to alter gene expression at a posttranscriptional level (Lackner and Bahler 2008) we hypothesized that Pum1 targets may be dynamic during cell stress. There is precedent for this condition-specific regulation by PUF proteins in yeast, where Puf3 differentially regulates stability of target messages when yeast are grown on different carbon sources (Foat, Houshmandi et al. 2005).

6.2 Results

6.2.1 Pum1 subcellular localization before and during oxidative stress

In order to determine the subcellular localization of Pum1 before and during stress we performed immunofluorescence analysis in HeLa cells, using a 45 minute

treatment with 0.5mM sodium arsenite as a stressor. We used the same antibody as was used in the RIP-Chip experiments along with a Cy-3 labeled secondary antibody to stain Pum1, along with either transfected fluorescently tagged stress granule and p-body markers (Figure 31) or in combination with staining for an endogenous stress granule or p-body marker (Figure 3). Results showed that Pum1 has a granular cytoplasmic staining pattern before stress, and during stress Pum1 colocalized with the endogenous stress granule marker HuR and several transfected stress granule markers: FAST, G3BP, and TIA-1 (Kedersha and Anderson 2007). However, Pum1 did not colocalize with either the endogenous PB marker Ddx6 or the transfected PB marker Dcp1a (Kedersha and Anderson 2007). It should be noted that Ddx6 is weakly present in Stress granules (Kedersha and Anderson 2007), as demonstrated by the weak signal overlapping with Pum1 staining.

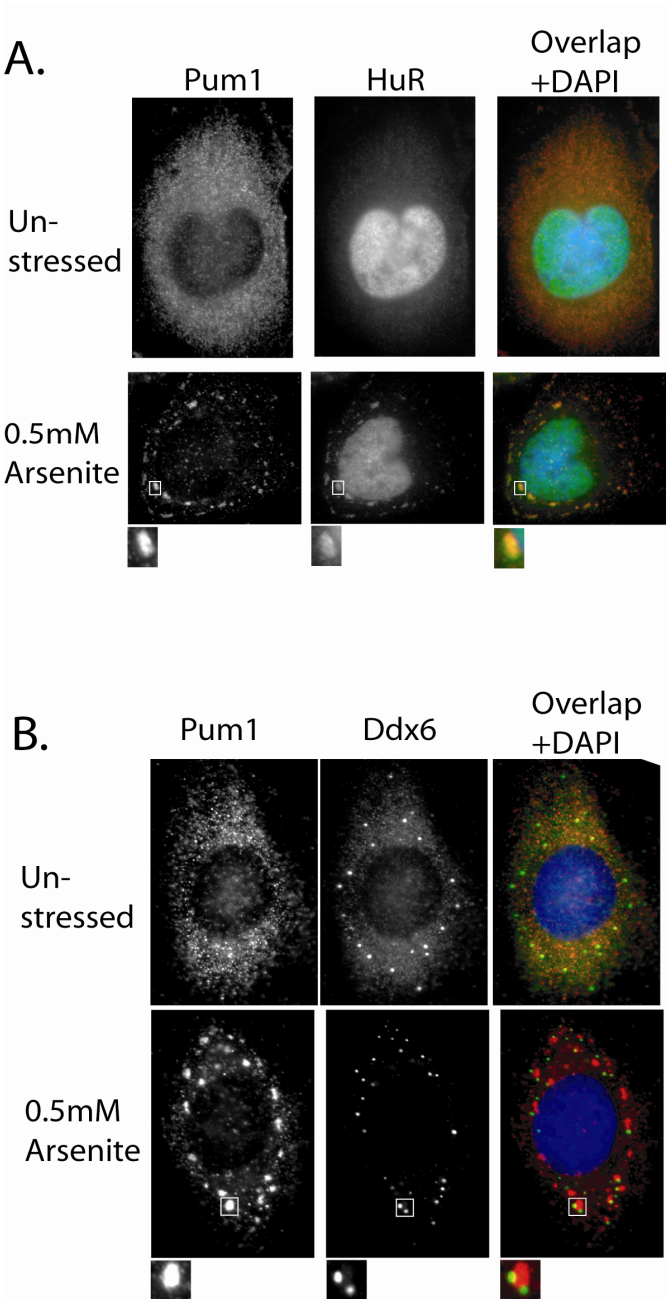


Figure 30: Subcellular dynamics of Pum1. Co-immunofluorescence of Pum1 (red) and HuR (A) or Ddx6 (B) (second panels and green) in unstressed and stressed HeLa cells is shown. DNA is stained with DAPI and is shown in blue. Magnified regions show a stress granule containing both Pum1 and HuR (A) and a Pum1 containing stress granule juxtaposed with Ddx6 containing p-bodies (B).

As an aside, initial co-staining with Pum1 and Ddx6 seemed to reveal that Pum1 was in both stress granules and p-bodies, which was in contrary to the stress granule only localization seen when visualizing Pum1 along with the Dcp1a construct. We determined that this discrepancy was caused by a lack of specificity of the anti-goat secondary antibody, which was binding either to the anit-Ddx6 primary antibody or the anti-rabbit secondary. We overcame this problem by first staining for Pum1 using both primary and secondary antibodies and then staining for Ddx6 after thorough washing. This strategy prevented non-specific binding and allowed us to visualize bona fide Pum1 protein, resulting in the conclusion that Pum1 is not in p-bodies but only stress granules, a conclusion supported by our previous result.

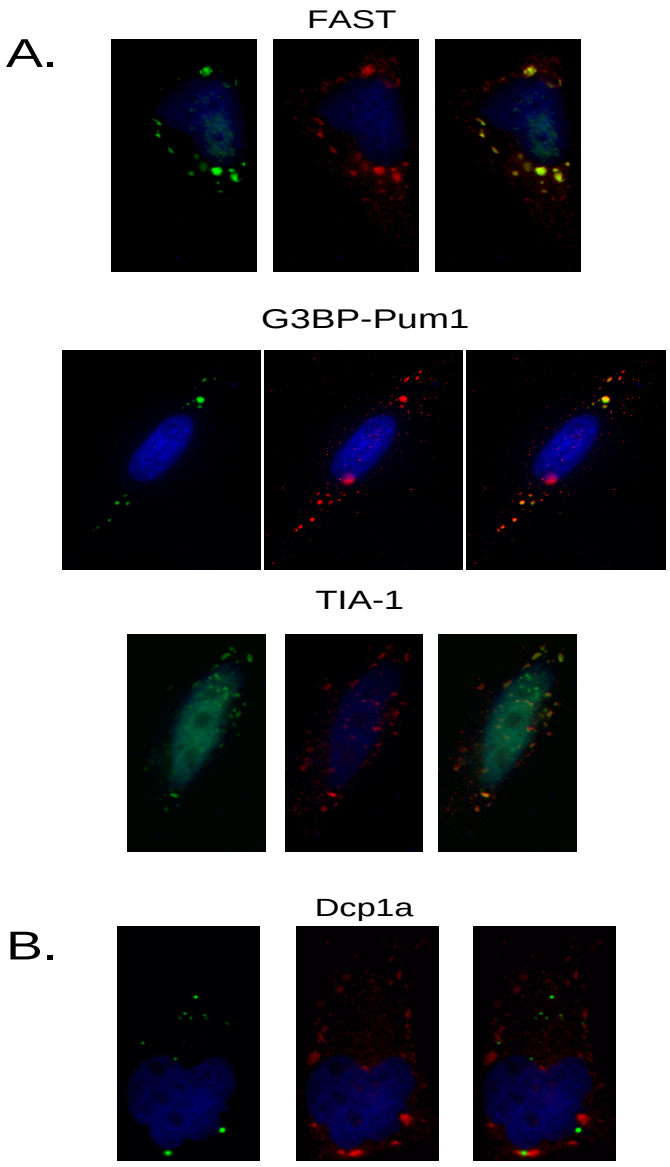


Figure 31: Confirmation of subcellular location of Pum1. Pum1 was visualized by immunofluorescence in cells that had been transfected with fluorescently tagged stress granule (A) or p-body (B) markers after arsenite treatment. In all cases Pum1 is red and the middle panels, the transfected marker is green and the first panels and DNA is blue. The third panels are the overlap of the first two.

6.2.2 Effect of Pum1 depletion on stress granule formation

As depletion of Pum2 protein has been shown to disrupt stress granule formation (Vessey, Vaccani et al. 2006), we also sought to determine whether Pum1 knockdown would result in the inability to form stress granules. As shown in Figure 32, knockdown of Pum1 did not appear to interfere with stress granule formation, however it did appear to increase the number of p-bodies present (not shown). We realized that this result could be due to non-specific effects of siRNA mediated protein knockdown. While increased p-bodies were not seen in the cells transfected with a non-targeting siRNA control, increased number of p-bodies were observed in cells treated with an siRNA specific to another RBP, HuR, indicating that the increased number of p-bodies was due to the siRNA mediated protein knockdown. This result is logical, as many proteins that are involved in siRNA mediated knockdown are also present in p-bodies, and p-bodies are hypothesized to be a location where siRNA and miRNA mediated decay of target mRNAs occurs (Anderson and Kedersha 2006). Pum1 not being necessary for formation of stress granules is also a logical result, as a larger number of RBPs have already been shown to be necessary for stress granule formation, and it is unlikely that every protein that is found in stress granules is necessary for their formation.

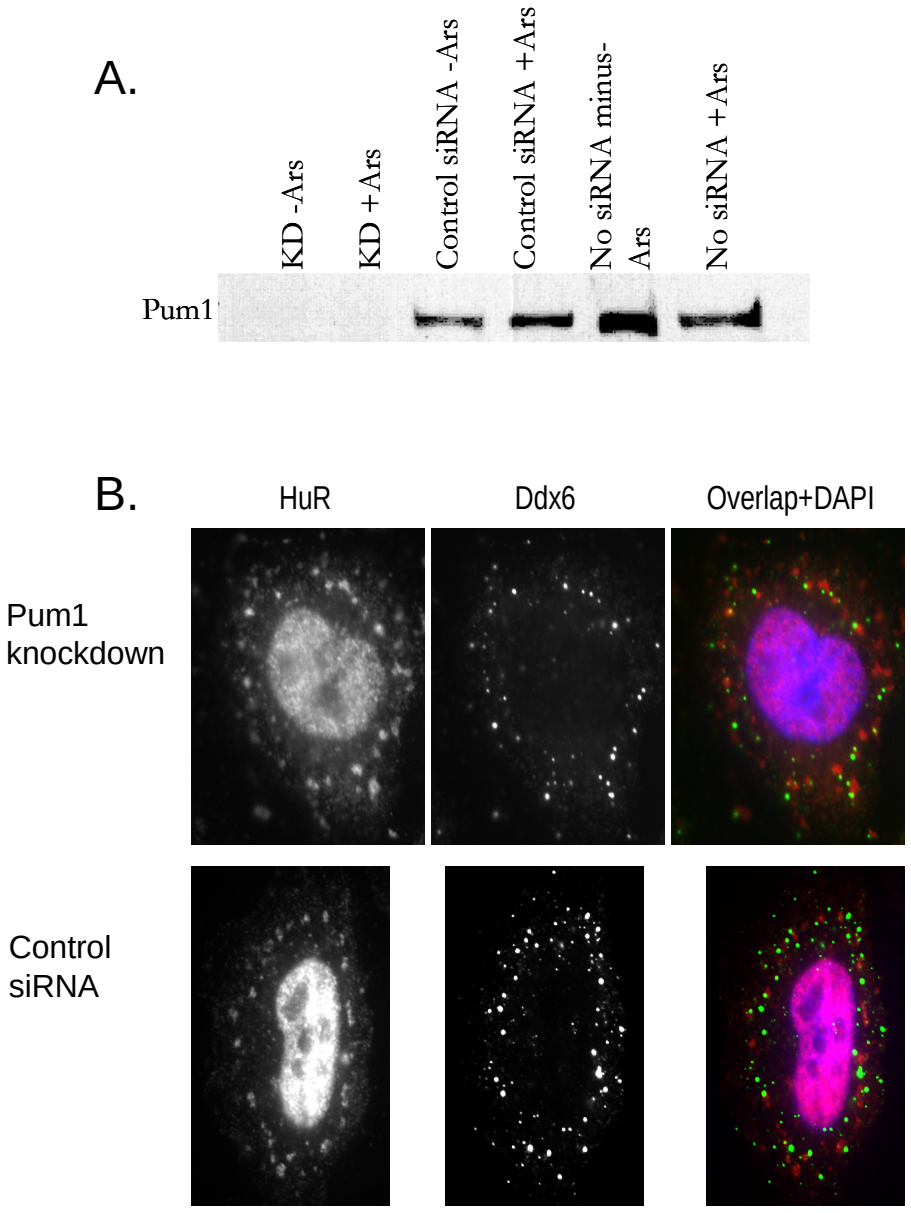


Figure 32: Pum1 knockdown does not prevent stress granule formation. Western blot confirming Pum1 knockdown is shown in (A), co-immunofluorescence of HuR (first panels and red) and Ddx6 (second panels and green) after arsenite treatment in Pum1 knockdown and control cells is shown. DNA is stained with DAPI and is shown in blue. Samples in (B) are represented by lanes 2 and 4 in (A).

6.2.3 Pum1 subcellular localization upon recovery from stress

We were also curious whether Pum1 containing stress granules would dissolve after removal of stress and indeed after addition of fresh media after stress Pum1 is completely re-localized (or de-localized) to a pattern resembling its pre-stress localization within 90 minutes. This result is consistent with localization of other stress granule components upon recovery from stress (Anderson and Kedersha 2008).

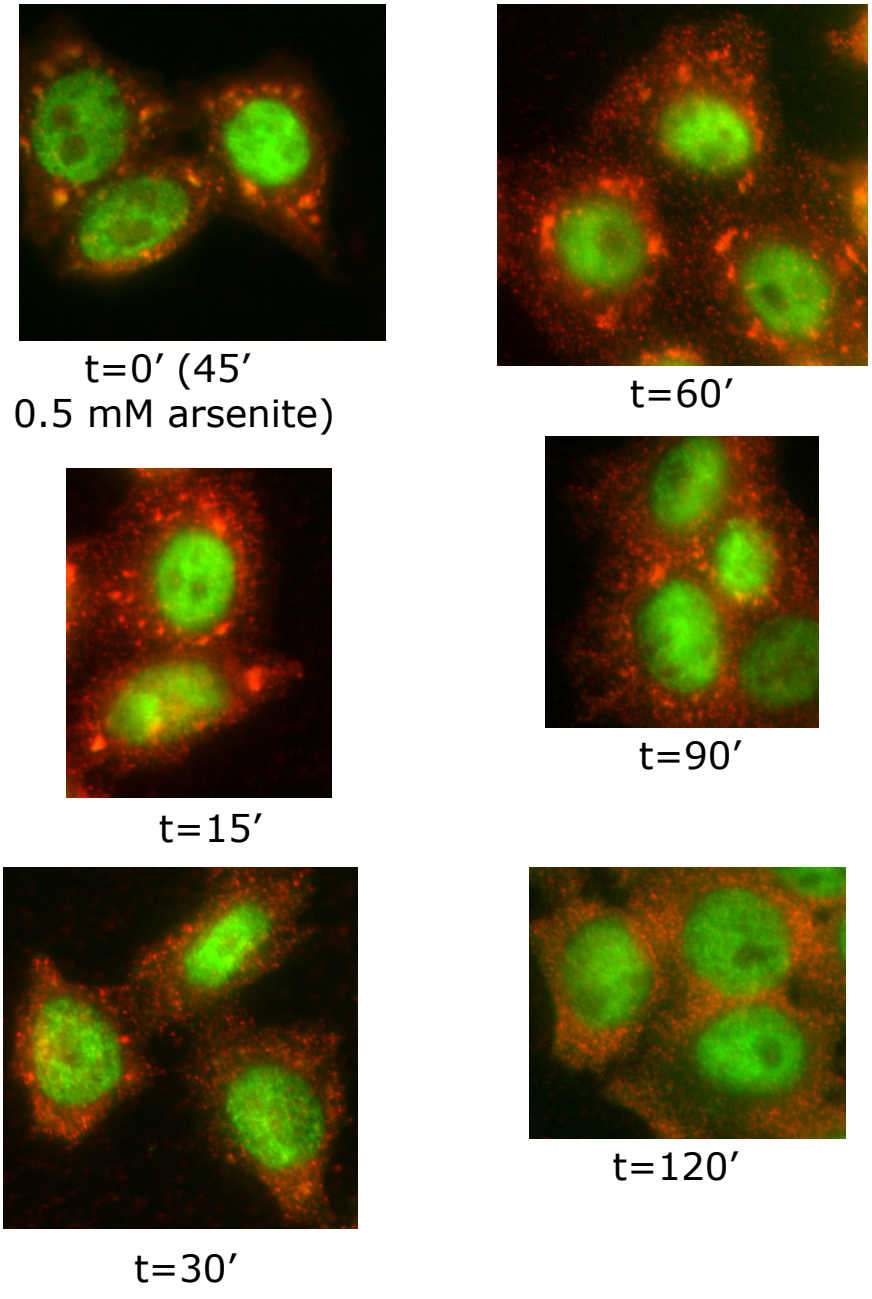
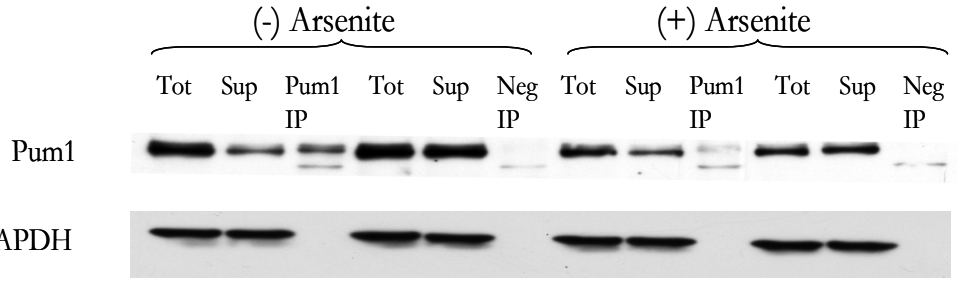


Figure 33: Pum1 stress granules disperse after removal of stress. Pum1 (red) and HuR (green) were visualized at indicated minutes after removal of arsenite and addition of normal media.

6.2.4 Pum1 mRNA targeting during stress

We measured the enrichment of Pum1 target mRNAs in Pum1 IPs versus negative IPs or total RNA before and during cell stress, with the results shown in Figure 34. Unexpectedly, while enrichments did change during stress, this change was simply a lessening of enrichment for all targets tested, which was likely due to less efficient recovery of Pum1 protein (Figure 34(A)). Thus, we concluded that targeting of Pum1 does not change during cell stress, although we cannot completely rule it out due to the relatively small number of mRNAs tested. We also cannot rule out the possibility that Pum1 indeed has lower affinity for its target mRNAs during cell stress, although this is unlikely given Pum1's role as a repressor and the observed lower efficiency of Pum1 recovery.

A.



B.

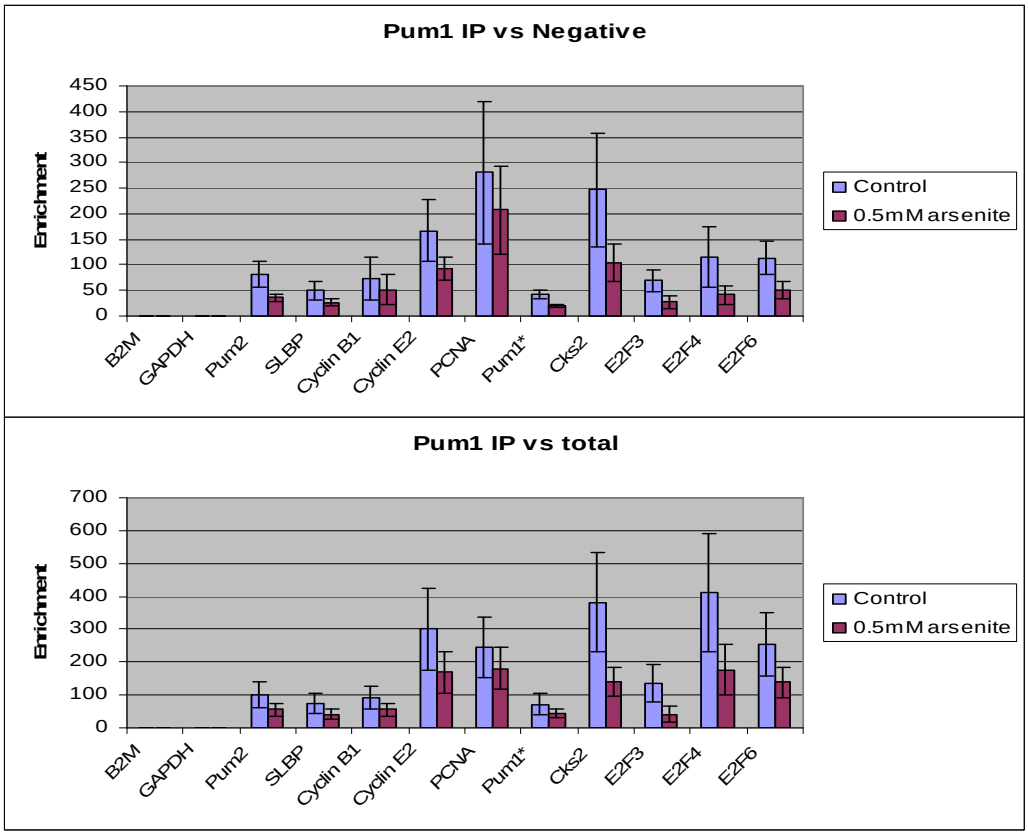


Figure 34: Pum1IP during stress. Pum1 protein was IPed from cells before and after addition of stress. Western blot in (A) shows Pum1 is recovered at slightly lower levels, RT-qPCR data in (B) demonstrates that Pum1 targets do not seem to be dynamic but instead are all recovered at lower levels after stress. qPCR data is normalized to B2M, bars represent SEM of three replicates.

6.3 Methods

HeLa CCL2 cells were used for immunofluorescence and were grown in DMEM supplemented with non-essential amino acids and 10% FBS.

All immunofluorescence was performed essentially as described previously (Kedersha and Anderson 2007). Anti HuR antibody was supernatant from hybridoma clone 3A2, anti-Ddx6 was from Bethyl labs. All fluorescent constructs were a gift from Dr. Nancy Kedersha.

Immunoprecipitation, RT-qPCR, and Western blots were performed as described previously in this document.

6.4 Discussion

Given the large number of RBPs that localize to stress granules, p-bodies, or both upon many types of stress (Kedersha, Stoecklin et al. 2005; Anderson and Kedersha 2006; Anderson and Kedersha 2008), and given the previous reported localization of Pum2 to stress granules (Vessey, Vaccani et al. 2006), it was not unexpected to find that Pum1 localizes to stress granules. Given the role we have demonstrated for Pum1 in decay of target mRNA and the proposed role of p-bodies as sites for mRNA decay (Anderson and Kedersha 2006), it seemed possible that Pum1 may also localize to p-bodies, although this was not the case. However, given that decapping enzymes are present in p-bodies (Anderson and Kedersha 2006) and Pum1 appears to act via deadenylation, which occurs before decapping, it may be that different steps of mRNA decay are happening in different locations. It is also possible that Pum1 mRNA targets are not degraded during stress but simply deadenylated and kept in a translationaly inactive state, allowing the cell to reuse them after the stress has been resolved without the need for abundant nascent transcription.

The finding that Pum1 targets were not dynamic under conditions of stress was unexpected as the targets of RBPs are often dynamic during changing cellular conditions (Morris, Mukherjee et al. 2009; Mukherjee, Lager et al. 2009). Given the lower recovery of Pum1 protein and associated mRNAs after stress, it is possible that the Pum1 protein

that is in stress granules is inaccessible, and thus the only Pum1 that is being assayed is that which is not found in stress granules, even after stress. It may also be possible that some stress granules are lost during the preparation of lysate, especially since it appears there is slightly less Pum1 protein in all samples after stress. Both of these situations may result in an observation of Pum1 associated mRNAs not being dynamic if the true dynamics directly related to whether the protein is in a stress granule.

7. Conclusions and Future Directions

RIP-Chip methodology has proven useful for studying a large number of RNA-binding proteins (Morris, Mukherjee et al. 2009), thus it was a logical approach for studying Pum1, a protein that very little was known about. Indeed, we were able to employ RIP-Chip to identify the genome-wide targets of Pum1. Among these targets we identified a number of putative RNA operons, many of which reflected the proposed ancestral role of PUF proteins in self-renewal of stem cells. We were also able to identify a sequence to which Pum1 likely binds in vivo, and in fact this sequence is identical to an RNA sequence that the RNA binding domain of Pum1 binds with high affinity (Wang, McLachlan et al. 2002).

Another study was published shortly after ours that employed similar RIP-Chip methodology to identify targets of Pum1 and Pum2 in HeLa cells, although the experimental specifics differed greatly. (Galgano, Forrer et al. 2008) Despite that, the results of both studies with relation to Pum1 were largely in agreement, with a high degree of overlap seen for target mRNAs and putative RNA operons and an identical Pum1 binding sequence identified (Galgano, Forrer et al. 2008).

The sequence bound by Pum1, UGUAHAUA, is also bound by Drosophila Pumilio (Gerber, Luschnig et al. 2006) and yeast Puf3 (Gerber, Herschlag et al. 2004), proteins that are fairly divergent from Pum1 outside of the RNA binding PUF domain, which is highly conserved (Wickens, Bernstein et al. 2002). Although the PUF domains

and sequences bound are almost identical, the mRNA targets of the three proteins are divergent (Gerber, Herschlag et al. 2004; Gerber, Luschnig et al. 2006; Galgano, Forrer et al. 2008; Morris, Mukherjee et al. 2008), suggesting that evolutionary rewiring has occurred through which the modules of the PUF domain and cognate binding sequence were conserved while the remainder of the proteins and the mRNA targets changed . This situation likely allowed the high affinity interaction between the two elements to be utilized for different purposes in species that have very different environmental demands (Mesarovic, Sreenath et al. 2004; Chan, Elemento et al. 2005; Gerber, Luschnig et al. 2006; Keene 2007; Halbeisen, Galgano et al. 2008).

It is believed that deadenylation is an important aspect of posttranscriptional regulation by acting as a major determinant of mRNA decay rates (Beelman and Parker 1995). Given that it is such a potentially important aspect of gene expression, current methods to study poly(A) tail lengths have serious drawbacks. One of the more commonly used methods, sometimes referred to as high resolution northern blotting, utilizes a digestion with RNAse H and a DNA oligo to separate a small portion of a 3' UTR from the rest of the mRNA, then a Northern blot is performed against this small portion of UTR (Salles, Richards et al. 1999). This method allows for accurate length determination of PATs, but requires a large amount of RNA, which may be prohibitive for genes that are expressed at low levels or situations where it is not possible to obtain a large amount of RNA, such as a human biopsy sample. In addition, analysis of multiple

mRNAs requires optimization of Northern blot conditions for each probe, which could potentially be very time consuming. Conversely, there are a few methods to determine PAT lengths based on PCR, which overcome the problems of needing a large sample size and difficulty of assaying multiple mRNAs (Salles, Richards et al. 1999). The drawback of these methods is that they are somewhat crude and not well suited for identifying small changes in PAT length. The GI-tailing method used here overcomes the issues associated with both types of assays. As it is PCR based it can easily be performed on a small sample or to assay multiple mRNA targets, while addition of GI tails allows priming for RT and PCR specifically at the end of the poly(A) tail, allowing resolution of small differences in PAT lengths. Indeed, while the PAT lengths of mRNAs assayed in this study sometimes differed only slightly, these changes were observed clearly and consistently. One caveat of all PAT assays, also suffered by GI-tailing, is that exact quantification is difficult because the result of the assay is in the form of a smear that represents a population of mRNAs with different PAT lengths. This caveat is why we avoided attempting to quantify PAT lengths in this study but simply made relative comparisons. Another caveat of this assay is that is uses a PCR step which sometimes requires high cycle numbers, which is especially prone to artifacts. This caveat, however, can be overcome by performing appropriate controls and using proper laboratory technique.

A number of proteins that interact with PUF proteins are known (Kraemer, Crittenden et al. 1999; Sonoda and Wharton 1999; Nakahata, Katsu et al. 2001; Sonoda and Wharton 2001; Hoek, Zanders et al. 2002; Jaruzelska, Kotecki et al. 2003; Urano, Fox et al. 2005; Ginter-Matuszewska, Spik et al. 2009), but at this point no in vivo binding partners of Pum1 have been identified. Results from this study indicate the Pum1 interacts with a deadenylase enzyme or complex, thus it will be of immediate interest to determine if this is the case. A portion of Pum1 protein was shown to interact with the deadenylase CNOT8 in vitro (Goldstrohm, Hook et al. 2006), although we were unable to reproduce this result. In addition to other deadenylases, it will also be interesting to see if other RNA binding proteins interact with Pum1. PUF proteins have been shown to bind to the RNA binding protein Nanos in other species (Kraemer, Crittenden et al. 1999; Sonoda and Wharton 1999; Nakahata, Katsu et al. 2001; Jaruzelska, Kotecki et al. 2003), thus it will be of interest to see if this interaction is conserved in human. Another advantage of identifying protein partners of Pum1 is that it may help to explain some unexpected results of the Pum1 analysis. For example, the differences seen in decay of target mRNAs may in part be explained by the specific composition of the Pum1 RNP, in particular the other proteins. Thus, by identifying potential binding partners of Pum1, it may be possible to specifically identify proteins bound by both Pum1 and another factor through a process of sequential immunoprecipitation. We have demonstrated that this type of experiment is feasible through sequential IP of two other

RBPs, Ago2 and HuR, as seen in Figure 35. A cell line was made that expressed a FLAG tagged version of Ago2, and an IP was performed against this FLAG tag. The beads were then washed and Ago2 protein released by elution with a molar excess of FLAG protein. This eluate was then used as input for an IP of HuR, and indeed we were able to IP HuR from the Ago2 enriched eluate. Recovery of RNA from this second IP should reveal mRNAs bound by both proteins. Using this type of strategy for Pum1, it should be possible to identify subsets of Pum1 targets that are bound by Pum1 and another protein, and discovery of these subsets may explain unexpected results of Pum1 analysis.

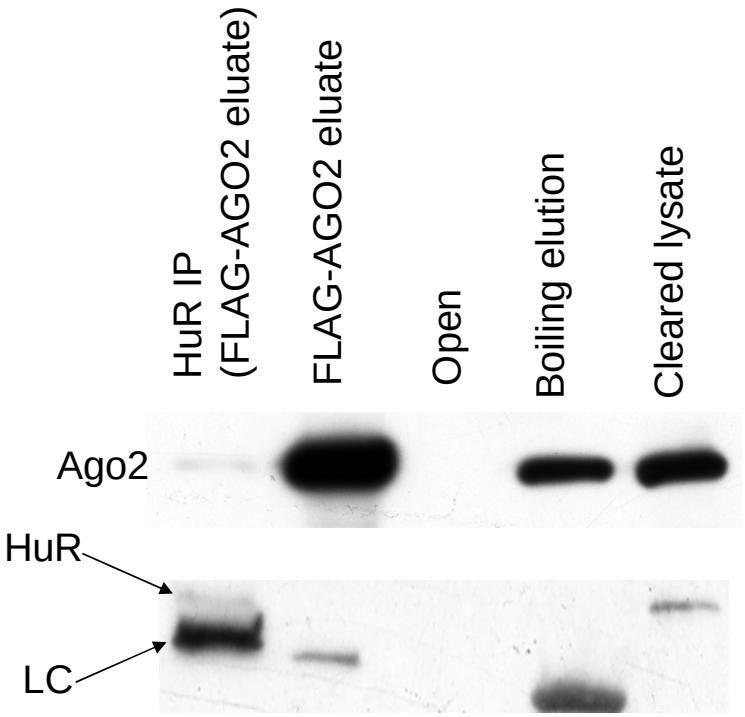


Figure 35: Sequential IP of Ago2 and HuR. Western blot demonstrating sequential Ago2-HuR IP as described in the text. Boiling elution sample is beads

from the Ago2 IP boiled in loading buffer after FLAG elution, cleared lysate is a sample of the input material for the Ago2 IP. LC=Light Chain

We were unable to demonstrate a function for Pum1 in normal cell cycle progression, although this result is likely due to the situation we were studying, which was normal cell cycle progression in exponentially growing HeLa cells. The likely ancestral function of PUF proteins is in stem cell maintenance and self-renewal, as well as helping to regulate the balance between proliferation and differentiation (Wickens, Bernstein et al. 2002). Thus, it will be important to study the function of Pum1 in a different environment, such as a stem cell line that can be induced to differentiate, in order to define its biological function. In fact, studies of PUF proteins in multi-cellular organisms are often performed in stem cells or undifferentiated cells (Lin and Spradling 1997; Forbes and Lehmann 1998; Crittenden, Eckmann et al. 2003; Jaruzelska, Kotecki et al. 2003), providing further evidence that Pum1 is likely to function in these types of cells. In addition, the zebrafish orthologue of a recently identified human protein with a PUF-like RNA-binding domain was shown to function in primordial germ cell migration and was expressed at early but not late stages of folliculogenesis in adult ovaries, providing further support for the function of PUF proteins in stem cells in vertebrae (Kuo, Wang et al. 2009).

While this document was being prepared, a ribonomics analysis of a C. elegans protein, FBF, was published (Kershner and Kimble 2010). This study found that there was a small but significant enrichment between mRNA targets of FBF and those of fly

Pumilio and human Pum1. Many of these conserved targets were involved in maintenance of stem cells, which led the authors to conclude that these conserved targets reflected the ancestral role of PUF proteins in maintenance of self-renewal of stem cells. The authors also found that FBF also likely functions as a "regulator of regulator" as it bound a number of mRNAs encoding regulatory proteins (Kershner and Kimble 2010).

Although we demonstrated that Pum1 enhances decay of a number of target mRNAs, these targets were somewhat "cherry-picked" and thus might not be a true representation of genome-wide Pum1 targets. It would be extremely interesting to combine a knockdown of Pum1 protein with a genome-wide determination of decay rates, which could then be compared back to Pum1's genome-wide targets. In fact, this type of comparison could make use of the probabilistic determination of Pum1 associated mRNAs, systematically determining how T-scores relate to effect on decay. Changes in decay rates could also be compared to a number of other 3' UTR characteristics, including number of Pum1 USERs, distance of the USER from the stop codon and poly-A tail, quality of the USER versus the consensus sequence, and length.

Pum1 is also likely to repress translation of target messages, either as a result of its effect on PAT length and decay or through an independent mechanism. Thus, it would also be interesting to study how global translation changes after Pum1 knockdown. Ideally, this data could be compared to decay and targeting data to

determine whether Pum1 has an effect on translation that is independent of its effect on decay.

One result observed consistently through this study was that Pum2 appears to be one of the strongest Pum1 targets, both in enrichment in Pum1 IPs and in effects seen after Pum1 perturbation. It is likely that the two proteins act redundantly, especially given the high degree of overlap in mRNA targets (Galgano, Forrer et al. 2008). Pum1 mRNA itself is also a target of Pum1 protein, and this condition may exist as a feedback loop to limit expression of Pum1 and Pum2. In addition, this situation may allow brief waves of gene expression of Pum1 targets. If repression of all target mRNAs by Pum1 is suddenly relieved through a posttranslational mechanism it is likely to result in rapid expression of target mRNAs. As Pum1 and Pum2 will also be more highly expressed as targets, they will begin to re-repress Pum1 target mRNAs, thus allowing for a brief wave of gene expression that is quickly resolved. In fact, these types of waves or bursts of gene expression are seen during development, where PUF proteins have been shown to play crucial roles (Spassov and Jurecic 2003).

Human stem cells are currently a highly studied area of biology due to their potential to treat numerous human diseases, and this popularity will likely increase as more reagents are made available. Understanding how gene expression is regulated in stem cells is crucial, and this goal cannot be achieved without understanding all aspects of regulation, including posttranscriptional. PUF family RNA binding proteins exist in

all eukaryotic organisms and have been proposed to have an ancestral function in self-renewal of stem cells due to their functions in stem cells in a number of species (Wickens, Bernstein et al. 2002). Thus, our understanding of posttranscriptional regulation in human stem cells would be greatly aided by understanding how PUF proteins perform their functions. The work presented here provides a strong foundation for future studies of human PUF proteins, which are likely crucial players in coordinating gene expression programs driving the proliferation and differentiation of human stem cells.

References

Anantharaman, V., E. V. Koonin, et al. (2002). "Comparative genomics and evolution of proteins involved in RNA metabolism." Nucleic Acids Res **30**(7): 1427-64.

Anderson, P. and N. Kedersha (2006). "RNA granules." J Cell Biol **172**(6): 803-8.

Anderson, P. and N. Kedersha (2008). "Stress granules: the Tao of RNA triage." Trends Biochem Sci **33**(3): 141-50.

Bachorik, J. L. and J. Kimble (2005). "Redundant control of the Caenorhabditis elegans sperm/oocyte switch by PUF-8 and FBF-1, two distinct PUF RNA-binding proteins." Proc Natl Acad Sci U S A **102**(31): 10893-7.

Baek, D., J. Villen, et al. (2008). "The impact of microRNAs on protein output." Nature **455**(7209): 64-71.

Bailey, T. L., N. Williams, et al. (2006). "MEME: discovering and analyzing DNA and protein sequence motifs." Nucleic Acids Res **34**(Web Server issue): W369-73.

Bartel, D. P. and C. Z. Chen (2004). "Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs." Nat Rev Genet **5**(5): 396-400.

Beelman, C. A. and R. Parker (1995). "Degradation of mRNA in eukaryotes." Cell **81**(2): 179-83.

Ben-Tabou de-Leon, S. and E. H. Davidson (2009). "Modeling the dynamics of transcriptional gene regulatory networks for animal development." Dev Biol **325**(2): 317-28.

Bernstein, D., B. Hook, et al. (2005). "Binding specificity and mRNA targets of a C. elegans PUF protein, FBF-1." Rna **11**(4): 447-58.

Blake, W. J., K. A. M, et al. (2003). "Noise in eukaryotic gene expression." Nature **422**(6932): 633-7.

Bradrick, S. S., E. Y. Dobrikova, et al. (2007). "Poly(A)-binding protein is differentially required for translation mediated by viral internal ribosome entry sites." Rna **13**(9): 1582-93.

Buness, A., W. Huber, et al. (2005). "arrayMagic: two-colour cDNA microarray quality control and preprocessing." Bioinformatics **21**(4): 554-6.

Cao, Q., K. Padmanabhan, et al. "Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition." Rna **16**(1): 221-7.

Chan, C. S., O. Elemento, et al. (2005). "Revealing posttranscriptional regulatory elements through network-level conservation." PLoS Comput Biol **1**(7): e69.

Crittenden, S. L., D. S. Bernstein, et al. (2002). "A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans." Nature **417**(6889): 660-3.

Crittenden, S. L., C. R. Eckmann, et al. (2003). "Regulation of the mitosis/meiosis decision in the Caenorhabditis elegans germline." Philos Trans R Soc Lond B Biol Sci **358**(1436): 1359-62.

DeGregori, J. and D. G. Johnson (2006). "Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis." Curr Mol Med **6**(7): 739-48.

Edwards, T. A., S. E. Pyle, et al. (2001). "Structure of Pumilio reveals similarity between RNA and peptide binding motifs." Cell **105**(2): 281-9.

Fiore, R., S. Khudayberdiev, et al. (2009). "Mef2-mediated transcription of the miR379-410 cluster regulates activity-dependent dendritogenesis by fine-tuning Pumilio2 protein levels." Embo J **28**(6): 697-710.

Firket, H. and P. Mahieu (1967). "[Synchronization of division induced in HeLa cells with an excess of thymidine. Studies on blockage of various stages of the cell cycle]." Exp Cell Res **45**(1): 11-22.

Foat, B. C., S. S. Houshmandi, et al. (2005). "Profiling condition-specific, genome-wide regulation of mRNA stability in yeast." Proc Natl Acad Sci U S A **102**(49): 17675-80.

Forbes, A. and R. Lehmann (1998). "Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells." Development **125**(4): 679-90.

Fox, M., J. Urano, et al. (2005). "Identification and characterization of RNA sequences to which human PUMILIO-2 (PUM2) and deleted in Azoospermia-like (DAZL) bind." Genomics **85**(1): 92-105.

Friedman, R. C., K. K. Farh, et al. (2009). "Most mammalian mRNAs are conserved targets of microRNAs." Genome Res **19**(1): 92-105.

Galgano, A., M. Forrer, et al. (2008). Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system. PLoS ONE. **3:** e3164.

Gao, F. B., C. C. Carson, et al. (1994). "Selection of a subset of mRNAs from combinatorial 3' untranslated region libraries using neuronal RNA-binding protein Hel-N1." Proc Natl Acad Sci U S A **91**(23): 11207-11.

Garcia-Rodriguez, L. J., A. C. Gay, et al. (2007). "Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast." J Cell Biol **176**(2): 197-207.

Gerber, A., D. Herschlag, et al. (2004). Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. Plos Biol. **2:** E79.

Gerber, A., S. Luschnig, et al. (2006). Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in Drosophila melanogaster. Proc Natl Acad Sci USA. **103:** 4487-92.

Ginter-Matuszewska, B., A. Spik, et al. (2009). "The SNARE-associated component SNAPIN binds PUMILIO2 and NANOS1 proteins in human male germ cells." Mol Hum Reprod **15**(3): 173-9.

Goldstrohm, A. C., B. A. Hook, et al. (2006). "PUF proteins bind Pop2p to regulate messenger RNAs." Nat Struct Mol Biol **13**(6): 533-9.

Goldstrohm, A. C., D. J. Seay, et al. (2007). "PUF protein-mediated deadenylation is catalyzed by Ccr4p." J Biol Chem **282**(1): 109-14.

Grigull, J., S. Mnaimneh, et al. (2004). "Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors." Mol Cell Biol **24**(12): 5534-47.

Gygi, S. P., Y. Rochon, et al. (1999). "Correlation between protein and mRNA abundance in yeast." Mol Cell Biol **19**(3): 1720-30.

Halbeisen, R. E., A. Galgano, et al. (2008). "Post-transcriptional gene regulation: from genome-wide studies to principles." Cell Mol Life Sci **65**(5): 798-813.

Hao, S. and D. Baltimore (2009). "The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules." Nat Immunol **10**(3): 281-8.

Hieronymus, H. and P. A. Silver (2004). "A systems view of mRNP biology." Genes Dev **18**(23): 2845-60.

Hoek, M., T. Zanders, et al. (2002). "Trypanosoma brucei expression-site-associated-gene-8 protein interacts with a Pumilio family protein." Mol Biochem Parasitol **120**(2): 269-83.

Hogan, D. J., D. P. Riordan, et al. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. Plos Biol. **6:** e255.

Ideker, T., V. Thorsson, et al. (2001). "Integrated genomic and proteomic analyses of a systematically perturbed metabolic network." Science **292**(5518): 929-34.

Jackson, J. S., Jr., S. S. Houshmandi, et al. (2004). "Recruitment of the Puf3 protein to its mRNA target for regulation of mRNA decay in yeast." Rna **10**(10): 1625-36.

Jaruzelska, J., M. Kotecki, et al. (2003). "Conservation of a Pumilio-Nanos complex from Drosophila germ plasm to human germ cells." Dev Genes Evol **213**(3): 120-6.

Kadyrova, L. Y., Y. Habara, et al. (2007). "Translational control of maternal Cyclin B mRNA by Nanos in the Drosophila germline." Development **134**(8): 1519-27.

Kaplan, J. C., A. Kahn, et al. (1992). "Illegitimate transcription: its use in the study of inherited disease." Hum Mutat **1**(5): 357-60.

Kedersha, N. and P. Anderson (2007). "Mammalian stress granules and processing bodies." Methods Enzymol **431**: 61-81.

Kedersha, N., G. Stoecklin, et al. (2005). "Stress granules and processing bodies are dynamically linked sites of mRNP remodeling." J Cell Biol **169**(6): 871-84.

Keene, J. D. (2001). "Ribonucleoprotein infrastructure regulating the flow of genetic information between the genome and the proteome." Proc Natl Acad Sci U S A **98**(13): 7018-24.

Keene, J. D. (2007). "RNA regulons: coordination of post-transcriptional events." Nat Rev Genet **8**(7): 533-43.

Keene, J. D., J. M. Komisarow, et al. (2006). "RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts." Nat Protoc **1**(1): 302-7.

Keene, J. D. and P. J. Lager (2005). "Post-transcriptional operons and regulons co-ordinating gene expression." Chromosome Res **13**(3): 327-37.

Keene, J. D. and S. A. Tenenbaum (2002). "Eukaryotic mRNPs may represent posttranscriptional operons." Mol Cell **9**(6): 1161-7.

Kershner, A. M. and J. Kimble (2010). "Genome-wide analysis of mRNA targets for Caenorhabditis elgans FBF, a conserved stem cell regulator." Proc Natl Acad Sci U S A **107**(8): 3936-41.

Kraemer, B., S. Crittenden, et al. (1999). "NANOS-3 and FBF proteins physically interact to control the sperm-oocyte switch in Caenorhabditis elegans." Curr Biol **9**(18): 1009-18.

Kuo, M. W., S. H. Wang, et al. (2009). "A novel puf-A gene predicted from evolutionary analysis is involved in the development of eyes and primordial germ-cells." PLoS One **4**(3): e4980.

Lackner, D. H. and J. Bahler (2008). "Translational control of gene expression from transcripts to transcriptomes." Int Rev Cell Mol Biol **271**: 199-251.

Lai, E. C. (2002). "Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation." Nat Genet **30**(4): 363-4.

Landthaler, M., D. Gaidatzis, et al. (2008). Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA.

Lee, M. H., B. Hook, et al. (2007). "Conserved regulation of MAP kinase expression by PUF RNA-binding proteins." PLoS Genet **3**(12): e233.

Lehmann, R. and C. Nusslein-Volhard (1987). "Involvement of the pumilio gene in the transport of an abdominal signal in the Drosophila embryo." Nature **329**(6135): 167-170.

Lewis, B. P., I. H. Shih, et al. (2003). "Prediction of mammalian microRNA targets." Cell **115**(7): 787-98.

Licatalosi, D. D., A. Mele, et al. (2008). "HITS-CLIP yields genome-wide insights into brain alternative RNA processing." Nature **456**(7221): 464-9.

Lim, L. P., N. C. Lau, et al. (2005). "Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs." Nature **433**(7027): 769-73.

Lin, H. and A. C. Spradling (1997). "A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary." Development **124**(12): 2463-76.

Lopez de Silanes, I., M. Zhan, et al. (2004). "Identification of a target RNA motif for RNA-binding protein HuR." Proc Natl Acad Sci U S A **101**(9): 2987-92.

Majoros, W. H. and U. Ohler (2007). "Spatial preferences of microRNA targets in 3' untranslated regions." BMC Genomics **8**: 152.

Maniatis, T. and R. Reed (2002). "An extensive network of coupling among gene expression machines." Nature **416**(6880): 499-506.

Mansfield, K. D. and J. D. Keene (2009). "The ribonome: a dominant force in co-ordinating gene expression." Biology of the Cell (101): 169-181.

Martin, G. and W. Keller (1998). "Tailing and 3'-end labeling of RNA with yeast poly(A) polymerase and various nucleotides." Rna **4**(2): 226-30.

Mazumder, B., P. Sampath, et al. (2003). "Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control." Cell **115**(2): 187-98.

Mee, C. J., E. C. Pym, et al. (2004). "Regulation of neuronal excitability through pumilio-dependent control of a sodium channel gene." J Neurosci **24**(40): 8695-703.

Menon, K. P., S. Andrews, et al. (2009). "The translational repressors Nanos and Pumilio have divergent effects on presynaptic terminal growth and postsynaptic glutamate receptor subunit composition." J Neurosci **29**(17): 5558-72.

Menon, K. P., S. Sanyal, et al. (2004). "The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E." Neuron **44**(4): 663-76.

Mesarovic, M. D., S. N. Sreenath, et al. (2004). "Search for organising principles: understanding in systems biology." Syst Biol (Stevenage) **1**(1): 19-27.

Mi, H., N. Guo, et al. (2007). "PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways." Nucleic Acids Res **35**(Database issue): D247-52.

Moore, M. J. (2005). "From birth to death: the complex lives of eukaryotic mRNAs." Science **309**(5740): 1514-8.

Morris, A. R., N. Mukherjee, et al. (2008). "Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets." Mol Cell Biol **28**(12): 4093-103.

Morris, A. R., N. Mukherjee, et al. (2009). "Systematic analysis of posttranscriptional gene expression." Wiley Interdisciplinary Reviews: Systems Biology and Medicine(Online).

Mukherjee, N., P. J. Lager, et al. (2009). "Coordinated posttranscriptional mRNA population dynamics during T-cell activation." Mol Syst Biol **5**: 288.

Mukhopadhyay, R., P. S. Ray, et al. (2008). "DAPK-ZIPK-L13a axis constitutes a negative-feedback module regulating inflammatory gene expression." Mol Cell **32**(3): 371-82.

Muraro, N. I., A. J. Weston, et al. (2008). "Pumilio binds para mRNA and requires Nanos and Brat to regulate sodium current in Drosophila motoneurons." J Neurosci **28**(9): 2099-109.

Murata, Y. and R. P. Wharton (1995). "Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in Drosophila embryos." Cell **80**(5): 747-56.

Nakahata, S., Y. Katsu, et al. (2001). "Biochemical identification of Xenopus Pumilio as a sequence-specific cyclin B1 mRNA-binding protein that physically interacts with a Nanos homolog, Xcat-2, and a cytoplasmic polyadenylation element-binding protein." J Biol Chem **276**(24): 20945-53.

Niehrs, C. and N. Pollet (1999). "Synexpression groups in eukaryotes." Nature **402**(6761): 483-7.

Olivas, W. and R. Parker (2000). "The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast." Embo J **19**(23): 6602-11.

Orphanides, G. and D. Reinberg (2002). "A unified theory of gene expression." Cell **108**(4): 439-51.

Padmanabhan, K. and J. D. Richter (2006). "Regulated Pumilio-2 binding controls RINGO/Spy mRNA translation and CPEB activation." Genes Dev **20**(2): 199-209.

Parisi, M. and H. Lin (2000). "Translational repression: a duet of Nanos and Pumilio." Curr Biol **10**(2): R81-3.

Pearson, K. (1894). "Contributions to the Mathematical Theory of Evolution." Philosophical Transactions of the Royal Society of London. A **185**: 71-110.

Penalva, L. O., M. D. Burdick, et al. (2004). RNA-binding proteins to assess gene expression states of co-cultivated cells in response to tumor cells. Mol Cancer. **3:** 24.

Perou, C. M., S. S. Jeffrey, et al. (1999). "Distinctive gene expression patterns in human mammary epithelial cells and breast cancers." Proc Natl Acad Sci U S A **96**(16): 9212-7.

Pullmann, R., Jr., H. H. Kim, et al. (2007). "Analysis of turnover and translation regulatory RNA-binding protein expression through binding to cognate mRNAs." Mol Cell Biol **27**(18): 6265-78.

Raj, A., C. S. Peskin, et al. (2006). "Stochastic mRNA synthesis in mammalian cells." PLoS Biol **4**(10): e309.

Rodriguez-Trelles, F., R. Tarrio, et al. (2005). "Is ectopic expression caused by deregulatory mutations or due to gene-regulation leaks with evolutionary potential?" Bioessays **27**(6): 592-601.

Salles, F. J., W. G. Richards, et al. (1999). "Assaying the polyadenylation state of mRNAs." Methods **17**(1): 38-45.

Salvetti, A., L. Rossi, et al. (2005). "DjPum, a homologue of Drosophila Pumilio, is essential to planarian stem cell maintenance." Development **132**(8): 1863-74.

Schweers, B. A., K. J. Walters, et al. (2002). "The Drosophila melanogaster translational repressor pumilio regulates neuronal excitability." Genetics **161**(3): 1177-85.

Selbach, M., B. Schwanhausser, et al. (2008). "Widespread changes in protein synthesis induced by microRNAs." Nature **455**(7209): 58-63.

Sonoda, J. and R. P. Wharton (1999). "Recruitment of Nanos to hunchback mRNA by Pumilio." Genes Dev **13**(20): 2704-12.

Sonoda, J. and R. P. Wharton (2001). "Drosophila Brain Tumor is a translational repressor." Genes Dev **15**(6): 762-73.

Souza, G. M., A. M. da Silva, et al. (1999). "Starvation promotes Dictyostelium development by relieving PufA inhibition of PKA translation through the YakA kinase pathway." Development **126**(14): 3263-74.

Spassov, D. S. and R. Jurecic (2002). "Cloning and comparative sequence analysis of PUM1 and PUM2 genes, human members of the Pumilio family of RNA-binding proteins." Gene **299**(1-2): 195-204.

Spassov, D. S. and R. Jurecic (2003). "The PUF family of RNA-binding proteins: does evolutionarily conserved structure equal conserved function?" IUBMB Life **55**(7): 359-66.

Spik, A., S. Oczkowski, et al. (2006). "Human fertility protein PUMILIO2 interacts in vitro with testis mRNA encoding Cdc42 effector 3 (CEP3)." Reprod Biol **6**(2): 103-13.

Subramanian, A., H. Kuehn, et al. (2007). "GSEA-P: a desktop application for Gene Set Enrichment Analysis." Bioinformatics **23**(23): 3251-3.

Tenenbaum, S. A., C. C. Carson, et al. (2000). Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. Proc Natl Acad Sci USA. **97:** 14085-90.

Tenenbaum, S. A., P. J. Lager, et al. (2002). "Ribonomics: identifying mRNA subsets in mRNP complexes using antibodies to RNA-binding proteins and genomic arrays." Methods **26**(2): 191-8.

Tennyson, C. N., H. J. Klamut, et al. (1995). "The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced." Nat Genet **9**(2): 184-90.

Ule, J., K. B. Jensen, et al. (2003). "CLIP identifies Nova-regulated RNA networks in the brain." Science **302**(5648): 1212-5.

Urano, J., M. S. Fox, et al. (2005). "Interaction of the conserved meiotic regulators, BOULE (BOL) and PUMILIO-2 (PUM2)." Mol Reprod Dev **71**(3): 290-8.

Vessey, J. P., L. Schoderboeck, et al. "Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function." Proc Natl Acad Sci U S A **107**(7): 3222-7.

Vessey, J. P., A. Vaccani, et al. (2006). "Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules." J Neurosci **26**(24): 6496-508.

Vyas, K., S. Chaudhuri, et al. (2009). "Genome-wide polysome profiling reveals an inflammation-responsive posttranscriptional operon in gamma interferon-activated monocytes." Mol Cell Biol **29**(2): 458-70.

Walser, C. B., G. Battu, et al. (2006). "Distinct roles of the Pumilio and FBF translational repressors during C. elegans vulval development." Development **133**(17): 3461-71.

Wang, X., J. McLachlan, et al. (2002). "Modular recognition of RNA by a human pumilio-homology domain." Cell **110**(4): 501-12.

Wang, X., P. D. Zamore, et al. (2001). "Crystal structure of a Pumilio homology domain." Mol Cell **7**(4): 855-65.

Wang, Y., C. L. Liu, et al. (2002). "Precision and functional specificity in mRNA decay." Proc Natl Acad Sci U S A **99**(9): 5860-5.

Wharton, R. P., J. Sonoda, et al. (1998). "The Pumilio RNA-binding domain is also a translational regulator." Mol Cell **1**(6): 863-72.

White, E. K., T. Moore-Jarrett, et al. (2001). "PUM2, a novel murine puf protein, and its consensus RNA-binding site." Rna **7**(12): 1855-66.

Whitfield, M. L., G. Sherlock, et al. (2002). "Identification of genes periodically expressed in the human cell cycle and their expression in tumors." Mol Biol Cell **13**(6): 1977-2000.

Wickens, M., D. S. Bernstein, et al. (2002). "A PUF family portrait: 3'UTR regulation as a way of life." Trends Genet **18**(3): 150-7.

Wilusz, C. J. and J. Wilusz (2007). "HuR-SIRT: the hairy world of posttranscriptional control." Mol Cell **25**(4): 485-7.

Xu, E. Y., R. Chang, et al. (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**(7): 912-21.

Yanai, I., J. O. Korbel, et al. (2006). "Similar gene expression profiles do not imply similar tissue functions." Trends Genet **22**(3): 132-8.

Zamore, P. D., J. R. Williamson, et al. (1997). "The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins." Rna **3**(12): 1421-33.

Zhang, B., M. Gallegos, et al. (1997). "A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line." Nature **390**(6659): 477-84.

Zhang, B., S. Kirov, et al. (2005). "WebGestalt: an integrated system for exploring gene sets in various biological contexts." Nucleic Acids Res **33**(Web Server issue): W741-8.

Zhang, H., J. Hu, et al. (2005). "PolyA_DB: a database for mammalian mRNA polyadenylation." Nucleic Acids Res **33**(Database issue): D116-20.

Biography

Adam Remy Morris was born August 27, 1982 in New Braunfels, Texas. Adam graduated from Texas A&M summa cum laude in May 2004 with a Bachelor of Science degree in Genetics.

Publications:

Daubner SC, McGinnis JT, Gardner M, Kroboth SL, Morris AR, Fitzpatrick PF. "A flexible loop in tyrosine hydroxylase controls coupling of amino acid hydroxylation to tetrahydropterin oxidation." J Mol Biol. 2006 Jun 2;359(2):299-307. Epub 2006 Mar 24

Morris, A. R., N. Mukherjee, et al. (2008). "Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets." Mol Cell Biol 28(12): 4093-103.

Morris, A. R., N. Mukherjee, et al. (2009). "Systematic analysis of posttranscriptional gene expression." Wiley Interdisciplinary Reviews: Systems Biology and Medicine(Online).

Awards:

"Best Talk", UPGG student retreat, May 2008.