Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Brittany Lynn Phillips Date

Tissue-specific regulation of the *Pabpn1* gene: Functional implications for muscular dystrophy

By

Brittany Lynn Phillips
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

Anita H. Corbett, Ph.D. Advisor	Grace K. Pavlath, Ph.D. Advisor
Yue Feng, Ph.D. Committee Member	Tamara Caspary, Ph.D. Committee Member
Daniel Reines, Ph.D. Committee Member	Kenneth Moberg, Ph.D. Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Tissue-specific regulation of the *Pabpn1* gene: Functional implications for muscular dystrophy

By

Brittany Lynn Phillips
B.S. University of Florida, 2013

Advisors: Anita Corbett, Ph.D. and Grace Pavlath, Ph.D.

An Abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Genetics and Molecular Biology
2018

Abstract

Tissue-specific regulation of the *Pabpn1* gene: Functional implications for muscular dystrophy

By Brittany Lynn Phillips

Fine-tuned spatio-temporal control of gene expression is often achieved through post-transcriptional regulatory mechanisms via a suite of RNA binding proteins. These RNA binding proteins are responsible for every aspect of RNA processing from 5' capping, splicing, and 3' end formation to eventual transcript degradation. Because many of these steps are necessary for proper RNA processing regardless of cell type, many genes encoding these RNA binding proteins are ubiquitously expressed. However, mutations in these genes encoding RNA binding proteins often cause tissue-specific diseases. Defining specific functions of RNA binding proteins in particular tissues will give insight into how mutations in these genes encoding RNA binding proteins cause tissue-specific disease. The nuclear poly(A) binding protein PABPN1, for example, plays critical roles in several RNA processing events including 3' end formation and polyadenylation. Although the *PABPN1* gene is ubiquitously expressed, a small expansion that leads to an increase in an alanine tract in the N-terminus of the PABPN1 protein causes the muscle-specific disease oculopharyngeal muscular dystrophy (OPMD). The work presented in this dissertation explores why specific skeletal muscles are affected by *PABPN1* mutations by characterizing novel mouse models of OPMD. These mouse models are critical tools allowing us to dissect the molecular phenotypes underlying OPMD pathology, including assessing the function of expanded PABPN1. Through this work we have determined that the alanine-expanded PABPN1 is only partially functional. As PABPN1 protein levels are very low in muscle compared to non-muscle tissues, a reduction in PABPN1 function could underlie pathology. We have identified multiple factors including the RNA binding protein HuR and microRNAs that regulate *Pabpn1* expression in muscle. Our work investigating the post-transcriptional mechanisms that contribute to low PABPN1 levels specifically in muscle provides insight into how PABPN1 expression is regulated and lays the groundwork for novel OPMD therapies seeking to raise PABPN1 levels in muscle. Together, our findings have elucidated how loss of PABPN1 function contributes to OPMD pathogenesis, and how muscle-specific regulation of PABPN1 levels could predispose this tissue to pathology.

Tissue-specific regulation of the *Pabpn1* gene: Functional implications for muscular dystrophy

By

Brittany Lynn Phillips
B.S., University of Florida, 2013

Advisors: Anita Corbett, Ph.D. and Grace Pavlath, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Genetics and Molecular Biology
2018

TABLE OF CONTENTS

Figures

3.8: HuR localization is more cytoplasmic in C2C12 myotubes than C2C12 myoblasts

4.1: The *Pabpn1* 3'UTR contains multiple conserved putative microRNA binding sites

4.2: microRNA-141-5p overexpression negatively regulates *Pabpn1* expression.

4.3: microRNA-532-3p positively regulates *Pabpn1* expression

4.4: microRNA-331-3p negatively regulates *Pabpn1* expression

5.1: The complement of *trans*-acting factors bound to the *Pabpn1* 3'UTR could influence *Pabpn1* expression in a tissue-specific manner.

5.2: Identification of loss of function phenotypes and mechanisms that contribute to low PABPN1 levels in muscle: Implications for OPMD treatment.

Tables

Chapter 1: Introduction

A portion of this chapter is adapted from the following published work:

Phillips B.L., Corbett A.H., Vest K.E. *PABPN1*.
Encyclopedia of Signaling Molecules, 2nd Edition. (2018) (3766-3772).

1.1 Scope and significance of this dissertation

RNA processing is a complex and highly regulated process that ensures proper spatial and temporal control of gene expression. Post-transcriptional control of gene expression is a common mechanism contributing to this cell or tissue-specific control of gene expression. Often this regulation occurs through 3'UTRs of transcripts as these regions contain *cis*-regulatory elements that interact with *trans*-acting factors. These *trans*-acting factors include microRNAs and RNA binding proteins that can influence protein levels of their target transcripts by regulating subcellular localization, RNA degradation, and/or translation. Many genes encoding RNA binding proteins are ubiquitously expressed, which is not surprising considering the basic roles these proteins play in RNA processing and gene expression. However, expression of these genes encoding RNA binding proteins is often dynamic across tissues and cell types. Furthermore, mutations in ubiquitously expressed genes encoding RNA binding proteins often cause tissue-specific pathology (1). My dissertation work has focused on one such RNA binding protein, PABPN1, which is ubiquitously expressed and plays critical roles in 3' end formation and polyadenylation. However, small expansion mutations in the *PABPN1* gene cause the muscle-specific disease OPMD. We have taken two approaches to understand why muscle-specific phenotypes arise when the *PABPN1* gene is mutated. First, we investigate phenotypes caused by endogenous levels of expanded PABPN1 in muscle, and whether reduced PABPN1 function contributes to these phenotypes. Second, we have identified post-transcriptional mechanisms that regulate wildtype *Pabpn1* transcript and protein levels in muscle with the goal of identifying pathways that could be

modulated to increase PABPN1 protein levels, which could be protective against OPMD pathology.

In Chapter 2, we have generated and characterized a novel mouse model which is the closest-available genocopy of OPMD patients (2). This novel OPMD model is critical to determine the molecular deficiencies caused by expanded PABPN1 as many previous studies have utilized either non-muscle cell lines such as HEK293 or HeLa cells, or mouse models with transgenic overexpression of expanded *Pabpn1*. Our novel mouse model, *Pabpn1*$^{A17/+}$, allows us to study multiple phenotypes, including histological and molecular deficiencies, without the confounding effects of transgenic *Pabpn1* overexpression. Analyzing PABPN1 in the context of muscle is critical as this is the only tissue typically affected in OPMD, which suggests that there are muscle-specific functions or requirements for PABPN1 which may not be recapitulated in other cells or tissues. In Chapter 2, we also describe a second novel mouse model of OPMD, a heterozygous *Pabpn1* knockout mouse, *Pabpn1*$^{+/\Delta}$. By comparing phenotypes observed in *Pabpn1*$^{A17/+}$ and *Pabpn1*$^{+/\Delta}$ mice, we can determine whether a partial loss of PABPN1 function contributes to OPMD pathology.

The work presented in Chapter 3 investigates the muscle-specific post-transcriptional mechanisms regulating *Pabpn1* transcript and protein levels, as previous work suggests that low PABPN1 protein levels in muscle could predispose this tissue to OPMD pathology. We identified a *cis*-regulatory element that interacts with the RNA binding protein HuR to negatively regulate *Pabpn1* transcript and protein levels in muscle. Importantly, this HuR-mediated regulatory mechanism is conserved *in vivo*, as

muscles from HuR knockout mice have significantly increased *Pabpn1* transcript and protein levels relative to muscles from control animals.

Chapter 4 focuses on further defining *trans*-acting factors that post-transcriptionally regulate *Pabpn1* transcript levels in muscle cells. We investigated three microRNAs, miR-141-5p, miR-331-3p, and miR-532-3p, which were bioinformatically predicted to bind motifs in the *Pabpn1* 3'UTR. Both overexpression and knockdown studies demonstrate that these microRNAs modulate *Pabpn1* transcript levels. Importantly, experiments using reporters show that this regulation occurs through the *Pabpn1* 3'UTR.

Together, the work presented here presents a model where muscle-specific post-transcriptional regulation of *Pabpn1* transcript contributes to low PABPN1 protein levels in muscle, which could predispose this tissue to OPMD pathology when the partially functional alanine-expanded PABPN1 is present. We have identified several regulatory pathways that could be targeted by therapies to raise PABPN1 protein levels in muscle in OPMD, which could be protective against pathology. Overall, our work demonstrates the importance of studying the roles of RNA binding proteins in tissues affected by RNA binding protein dysfunction, as this may lead to a better understanding of and appreciation for tissue-specific requirements for post-transcriptional processing and proper gene expression.

1.2 Co-transcriptional and post-transcriptional regulation of gene expression

The central dogma of molecular biology represents how genetic information encoded in DNA is transcribed into the messenger RNA molecule, which is then translated to generate proteins that can function in a myriad of cellular processes. However, this simplified gene expression cascade is, in practice, highly complex with many regulatory mechanisms coalescing to modulate gene expression. This ability to specifically regulate gene expression ensures that cells are able to fulfill dynamic cellular requirements, a subset of which are cell type specific. Regulation at the RNA level, which can occur both co- and post-transcriptionally, fine-tunes multiple aspects of gene expression. For example, regulatory elements present at 3' ends of transcripts modulate subcellular localization, RNA stability, and translation, which together regulate the accessibility and/or amount of a particular RNA, ultimately dictating levels of the encoded protein. Overall, regulation at the level of RNA provides spatial and temporal control of gene expression, allowing the cell to tightly regulate the production of gene products.

1.2.1 RNA processing: An overview

RNA processing is necessary for proper spatial and temporal control of gene expression. Throughout the lifetime of an RNA, the RNA molecule is never "naked," but is covered by RNA binding proteins that enact these processing steps. Together, these RNA-RNA binding protein complexes form messenger ribonucleoproteins (mRNPs). The

RNA binding proteins associated with the mRNP are dynamic and dictate RNA processing steps (Figure 1.1). Briefly, the first RNA processing step is the co-transcriptional addition of a 5' N7-methylated guanosine, termed the 5' cap, that protects the transcript from 5' to 3' exonucleases and aids in recruiting additional RNA processing machinery to the transcript (3). The majority of splicing also occurs co-transcriptionally, with a large ribonucleoprotein, the spliceosome, coordinating the formation of a lariat structure between the 5' splice site and the branch point sequence, which allows eventual exon ligation between the 5' and 3' splice sites (4). In order to form the 3' end of the transcript, polyadenylation factors associate with the C-terminal domain (CTD) of the large subunit of RNA polymerase II to facilitate recognition of the polyadenylation signal (PAS) once it is transcribed. Polyadenylation machinery, including poly(A) polymerase, cleavage stimulation factor, and the nuclear poly(A) binding protein PABPN1 then commence with 3' end formation and polyadenylation reactions (5,6). Properly processed mRNAs are then exported from the nucleus into the cytoplasm, where they are translated and eventually degraded (7).

1.2.2 The CTD of the largest subunit of RNA polymerase II mediates co-transcriptional processing

Some RNA processing events, including 5' capping and 3' end formation, occur co-transcriptionally through physical association of factors with RNA polymerase II, which transcribes mRNAs that are translated to produce proteins. The evolutionarily-conserved CTD of RNA polymerase II (8) contains heptad repeats (Y1S2P3T4S5P6S7)

(9) that serve as a binding scaffold for various proteins including RNA processing factors. CTD phosphorylation is dynamic throughout transcription and influences what factors are bound at a given point in transcription (Figure 1.2). This phosphorylation-mediated protein association ensures that co-transcriptional processing events, including 5' capping and 3' end formation, occur during the appropriate stage of transcription.

While Serine 5 is phosphorylated (Ser5P) during transcriptional initiation, elongating RNA polymerase II is thought to be phosphorylated at both Ser5 and Ser2 (Ser2P), perhaps as a transitional phase to exclusive phosphorylation at Ser2 during transcriptional termination (9-11). Therefore, protein complexes that bind the CTD in a phosphorylation-dependent manner can recruit RNA processing factors during the appropriate stage of transcription. For example, the 5' capping enzyme binds to Ser5P of the CTD (12), resulting in 5' cap (7-methyl G5'ppp5'N) addition after ~25 nucleotides have emerged from the RNA polymerase II exit channel (13). This specificity of Ser5P during transcription initiation ensures that capping machinery is recruited when the 5' end of the transcript is poised for the capping reaction.

Coupling of transcription and 3' end formation ensures that cleavage and polyadenylation factors are poised to recognize the PAS and commence with cleavage and polyadenylate (poly[A]) tail addition. Both cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulation factor (CstF) interact with the CTD of RNA polymerase II during transcriptional elongation to prepare for transcriptional termination (5,6) and 3' end processing when the PAS emerges from the RNA polymerase II exit channel. In fact, the PAS is important for transcriptional termination as mutations disrupting PAS sequences cause termination defects (14,15), demonstrating that

transcription termination and 3' end formation are tightly coupled. Overall, modulation of co-transcriptional processing by the CTD ensures that processing machinery are poised during the appropriate stage of transcription.

1.2.3 The 3'UTR contains *cis*-regulatory elements that interact with *trans*-acting factors

After RNAs are co-transcriptionally capped, spliced, and polyadenylated, post-transcriptional regulation ensures that expression is spatially and temporally controlled. The 3' untranslated region (3'UTR) of a transcript downstream of the stop codon mediates many post-transcriptional regulatory steps. 3'UTRs can contain *cis*-regulatory elements that interact with *trans*-acting factors to influence processing steps such as mRNA export, RNA stability, translation and protein-protein interactions that seed protein complex formation (16).

3'UTRs are evolutionarily conserved from bacteria to humans, and correlations have been observed between 3'UTR length and genome complexity. The median length of bacterial 3'UTRs is ~60 nucleotides whereas the median length of human 3'UTRs is ~1200 nucleotides (16). This finding suggests that longer 3'UTRs are important for coordinating complex gene expression requirements in higher organisms, such as tissue-specific regulation of expression.

Although 3'UTR sequences as a whole are not as highly conserved as coding region sequences, microRNA or RNA binding protein binding sites within 3'UTRs are more highly conserved than regions within 3'UTRs that do not contain these regulatory

sequences (17,18). By binding to 3'UTRs, *trans*-acting factors modulate multiple aspects of post-transcriptional gene expression from subcellular localization to translation. These *trans*-acting factors include RNA binding proteins and microRNAs that regulate critical aspects of gene expression at the RNA level such as subcellular localization, RNA stability, and translation.

3'UTRs can contain motifs recognized by RNA binding proteins such as AU-rich elements (AREs), which are a well-studied class of *cis*-regulatory elements that influence transcript stability. Three classes of AREs have been identified: Class I, which contains two AUUUA pentamers; Class II, which contains multiple clustered AUUUA pentamers (19); and Class III, which contains degenerate AU- or U-rich tracts (19). A subset of RNA binding proteins, termed ARE-binding proteins, recognize these ARE motifs and modulate target stability (20). These ARE-binding proteins can destabilize target transcripts by recruiting degradation machinery, or stabilize target transcripts potentially by preventing binding of destabilizing factors. Overall, AREs are relatively rare and only occur in ~8% of transcripts, and many transcripts containing AREs are unstable and require precise spatial and temporal control of gene expression (20).

Trans-acting factors recognize these *cis*-regulatory elements within 3'UTRs to post-transcriptionally modulate gene expression. These *trans*-acting factors, such as microRNAs and RNA binding proteins, confer spatial and temporal control of gene expression by regulating critical events such as RNA stability and translation. microRNAs are ~22 nucleotide non-coding RNAs that post-transcriptionally regulate expression of target transcripts most commonly through binding 3'UTRs. A "seed" sequence from positions 2 to 7 of the microRNA binds target transcripts through Watson-

Crick pairing (21), and this seed is essential for target recognition (22). After target recognition, the microRNA binding complex either degrades target transcripts or represses translation of these transcripts, thereby reducing levels of the encoded protein.

microRNA-mediated regulation of gene expression is widespread, with approximately 60% of mammalian mRNAs containing conserved microRNA binding sites (17). Therefore, microRNA-mediated regulation of diverse pathways from stress response to muscle differentiation is not surprising (23,24). As further evidence of the critical functions of microRNAs, aberrant microRNA expression is associated with many human diseases including cancer (25)and neurodegeneration (26,27). Therefore, microRNA-mediated regulation of expression is necessary to maintain proper function at both the cellular and tissue levels.

RNA binding proteins also bind specific motifs within 3'UTRs to modulate RNA processing events including mRNA export, mRNA stability, and translation. The RNA export factor Aly/REF associates with mRNAs co-transcriptionally and, through association with various export factors, modulates export from the nucleus into the cytoplasm (28). Once properly processed, mRNAs are exported to the cytoplasm where they can be translated and eventually degraded. RNA binding proteins, such as FMRP, can also modulate translation dynamics. FMRP suppresses translation of target transcripts (29), and structural studies show that binding of FMRP to the 80S ribosomal subunit could suppress translation by blocking access of tRNAs and translation elongation factors to the ribosome (30).

Modulating RNA degradation is one method the cell uses to regulate the levels of mRNAs that are available for translation. Human Antigen R (HuR) is a well-

characterized ARE-binding protein that stabilizes target transcripts perhaps by suppressing recruitment of degradation machinery. However, recent work has discovered that HuR can also destabilize or negatively regulate a subset of transcripts (31) demonstrating that HuR-mediated regulation of gene expression is complex and likely depends on the context of the target transcript and cellular environment, which will be discussed further in Chapter 3.

1.2.4 RNA binding proteins and target recognition

The interplay between these *cis*-elements and *trans*-acting factors, including the complement of regulatory factors present under particular cellular conditions, modulates gene expression. As previously described, RNA binding proteins are *trans*-acting factors that decorate mRNAs, forming mRNPs. However, these mRNPs are extensively remodeled during the lifetime of an mRNA, and some of these changes in mRNP composition reflect particular stages of RNA processing, such as nuclear export or translation (32). Furthermore, many components of these mRNPs are involved in multiple aspects of RNA processing and thus remain associated throughout the lifetime of the RNA. For instance, the poly(A) binding protein PABPN1, a component of the polyadenylation machinery, is involved in 3' end formation and polyadenylation but has been implicated in later steps of mRNA processing including nuclear export and the pioneer round of translation. Therefore, PABPN1 remains bound to transcripts in the cytoplasm before being replaced by the conventional cytoplasmic poly(A) binding protein, PABPC1 (33).

Defining the factors that regulate RNA binding protein recognition of particular mRNA targets is critical as this regulation is necessary to fulfill cellular and tissue-specific requirements for gene expression. RNA binding protein-mRNA interactions are mediated by domains within the RNA binding protein that recognize and bind sequence motifs within the target transcript. RNA binding domains, such as the RNA recognition motif (RRM) and the hnRNPK-homology (KH) domain, recognize single-stranded RNAs (34), although some RNA binding proteins bind preferentially to double-stranded RNAs. Crosslinking-immunoprecipitation followed by sequencing (CLIP-seq) studies have identified transcriptome-wide RNA targets of many RNA binding proteins, and motif analysis of these CLIP-seq datasets can determine whether RNA binding proteins recognize prefer specific binding motifs.

However, some RNA binding proteins recognize degenerate motifs, and the presence of a specific motif does not guarantee that the RNA binding protein will bind that motif under any or all cellular conditions. RNA binding protein recognition of and binding to target transcripts is often dynamic and dependent on spatial and temporal cues based on cellular demands. Furthermore, the presence of RNA modifications, secondary structure (35), or *trans*-acting factors such as microRNAs or other RNA binding proteins could influence motif recognition/binding by a particular RNA binding protein, thereby modulating expression.

Differential expression of genes encoding RNA binding proteins also influences expression of target transcripts. The majority of RNA binding proteins are ubiquitously expressed, which is not surprising considering the basic functions in RNA processing these proteins fulfill. However, some RNA binding proteins are tissue-specific or have

tissue-specific expression profiles, suggesting that requirements for RNA processing machinery are complex and depend on cellular context.

1.3 PABPN1: A ubiquitous RNA binding protein that causes tissue-specific disease

Mutations occurring in ubiquitously-expressed genes encoding RNA binding proteins often cause tissue-specific phenotypes (1). Nervous and muscle tissue are disproportionately affected by mutations in genes encoding RNA binding proteins (1,36), suggesting that these complex tissues have strict requirements for proper control of gene expression, and that RNA binding proteins could have tissue-specific functions. Small expansion mutations in the ubiquitously-expressed *PABPN1* gene, which encodes the poly(A) binding protein PABPN1, cause the muscle-specific disease oculopharyngeal muscular dystrophy (OPMD). Understanding the functions of and requirements for PABPN1 in muscle is necessary to determine why this tissue is specifically affected in OPMD.

1.3.1 OPMD: A rare, late-onset muscle disease

OPMD is an autosomal dominant, muscle-specific disease in which patients experience weakness in specific skeletal muscles including those of the eyelids, pharynx, and proximal limbs that cause eyelid drooping, difficulty swallowing, and loss of mobility (33). These symptoms greatly reduce quality of life with many patients experiencing malnutrition, aspiration pneumonia, and requiring wheelchair assistance. While OPMD is rare in Western populations with ~1:100,000 people affected, disease

prevalence is higher in particular bottleneck populations such as French Canadian (1:1000) and Bukhara Jewish (1:600) populations (37).

OPMD patients harbor small GCN expansion mutations in the *PABPN1* gene that expand a 10 alanine tract at the N-terminus of the PABPN1 protein to 11-18 alanines in OPMD patients (Figure 1.1A) (38). The expansion to 11 alanines is inherited in an autosomal recessive manner, but expansions as small as 12 alanines cause autosomal dominant disease (38). Thus, a single copy of *PABPN1* encoding a mere two additional alanines in a region of the protein with unknown function can cause disease. The molecular consequences of these small expansion mutations are unknown. Furthermore, OPMD is a late onset disease with the majority of patients presenting symptoms in the sixth decade of life. Why small expansion mutations in the ubiquitously-expressed *PABPN1* gene cause a late-onset disease affecting a subset of muscles is not understood.

As mutations in many ubiquitously-expressed genes encoding RNA binding proteins cause tissue-specific disease, it is clear that requirements for RNA processing events are dynamic across tissues. Therefore, studies investigating PABPN1 function in muscle and in the context of OPMD serve as a model for the importance of studying RNA binding proteins in diverse cellular and tissue contexts, as these proteins likely have dynamic regulatory functions depending on tissue or cell type-specific requirements.

1.3.2 PABPN1 functions: An overview

Understanding the basic functions of PABPN1 in a variety of tissues, including muscle, is critical to determine why muscle specifically is affected by a small alanine

expansion in the PABPN1 protein. PABPN1 plays multiple roles in polyadenylation from PAS selection and proper 3' end formation to modulating poly(A)-dependent quality control mechanisms (2,39-44). Despite ubiquitous expression of *PABPN1*, small expansion mutations in the *PABPN1* gene cause the skeletal muscle-specific disease OPMD (38). The molecular mechanisms that dictate pathology in particular skeletal muscles when PABPN1 is altered, while other muscles and tissues remain unaffected, is unknown.

The PABPN1 protein plays a variety of roles in RNA processing that are summarized in Figure 1.3B. The best characterized roles for PABPN1 involve polyadenylation, which will be described in detail below. Briefly, PABPN1 is critical for 3'-end formation, including poly(A) site (PAS) selection and subsequent polyadenylation (42,45-47), which are critical for fine-tuning gene expression. Polyadenylation contributes to transcript stability as RNAs that are not properly polyadenylated are targeted for degradation. Furthermore, PAS utilization can affect gene expression by influencing the length of the 3' untranslated region (UTR) which can alter regulation and expression of these transcript variants.

These roles for PABPN1 in polyadenylation are linked to characterized functional domains in the PABPN1 protein (Figure 1.3A) (33). The single RNA recognition motif (RRM) binds polyadenosine RNA and is involved in PABPN1 oligomerization during polyadenylation while the coiled-coil domain (CCD) mediates poly(A) polymerase (PAP) stimulation. The PABPN1 C-terminal domain also mediates poly(A) RNA binding and PABPN1 oligomerization, and contains a nuclear localization signal. At the N-terminus of the PABPN1 protein is a stretch of ten alanine residues immediately following the

initial methionine residue. In patients with OPMD, this alanine stretch is expanded to 11-18 alanine residues (38).

1.3.3 PABPN1 and 3' end formation

Polyadenylation is a complicated process that requires orchestration of several large multi-protein complexes. After RNA polymerase II transcribes the PAS, AAUAAA, PAP and cleavage and polyadenylation specificity factors (CPSF) initiate cleavage and polyadenylation. CPSF weakly binds the PAS in the nascent transcript (48), and this binding is strengthened by binding of the cleavage stimulation factors (CstF) to a GU-rich downstream sequence element (DSE) (49), resulting in endonucleolytic cleavage and a free 3' OH where the nascent poly(A) tail is distributively added by PAP. PABPN1 binding to the nascent poly(A) tail stimulates PAP, resulting in processive polyadenylation that terminates after ~250 adenosines have been added (46,47). In addition to stimulating polyadenylation, PABPN1 is thought to act as a molecular ruler to ensure that the proper number of adenosine residues are added to 3' ends by affecting the interaction between PAP and CPSF (45). In cultured primary muscle cells, *Pabpn1* knockdown results in global poly(A) tail shortening and nuclear retention of polyadenylated RNAs (50). PABPN1, therefore, plays a critical role in global modulation of poly(A) tail length, which in turn affects nuclear export efficiency.

Approximately 50-70% of mammalian transcripts contain multiple poly(A) sites, which are predominantly located within the 3'UTR (51). Choice of PAS determines 3'UTR length and therefore inclusion or exclusion of *cis*-regulatory elements such as

microRNA binding sites and motifs recognized by RNA binding proteins. Binding of microRNAs or RNA binding proteins regulates gene expression by repressing translation, recruiting or excluding RNA decay machinery, and influencing RNA localization. Studies in cultured cells show that *Pabpn1* knockdown results in a global shift in PAS usage leading to 3'UTR shortening for a majority of transcripts (42,44). Similar changes in PAS usage with *Pabpn1* knockdown were found for a subset of candidate genes in a muscle cell line (41). The mechanism by which PABPN1 regulates PAS usage has not been determined though the prevailing model suggests that PABPN1 binds to and masks the weaker proximal PAS from cleavage and polyadenylation machinery, thus enhancing distal PAS recognition and utilization (42).

1.3.4 Additional roles for PABPN1 in RNA processing

In addition to its canonical role in polyadenylation, PABPN1 influences splicing and translation. Although the majority of splicing occurs co-transcriptionally, some 3' introns are not spliced before polyadenylation is completed. Studies using reporter constructs as well as endogenous targets in the context of *Pabpn1* knockdown showed that PABPN1 influences splicing of terminal introns (52). Further mutational analysis revealed that the ability of PABPN1 to bind RNA and stimulate PAP is critical for modulating splicing at the 3' end of transcripts (52).

While the steady-state localization of PABPN1 protein is nuclear, the PABPN1 protein shuttles between the nucleus and the cytoplasm (33). Most studies have focused on nuclear roles of PABPN1, but some work has investigated potential roles in

translation. PABPN1 has been identified as part of a protein complex containing 5' cap proteins that is involved in mRNA quality control and the pioneer round of translation (53). Pab2, the PABPN1 ortholog in *S. pombe*, associates with polysomes (54), suggesting that PABPN1 may play a larger role in translation than previously appreciated. However, the specific mechanism by which PABPN1 participates in regulation of translation remains unknown.

Much of PABPN1 function has been described in the context of mRNAs, but PABPN1 also regulates a variety of non-coding RNAs by recruiting the RNA exosome machinery to targets for degradation. PABPN1-mediated hyperadenylation is required for rapid RNA exosome-mediated degradation of intronless, polyadenylated reporter RNAs (55). While *PABPN1* depletion has only mild effects on steady-state levels of mRNAs in HeLa cells, stability of a subset of polyadenylated long non-coding RNAs (lncRNAs) increases in the absence of PABPN1, likely due to loss of RNA exosome-mediated degradation (39). Polyadenylation is required for PABPN1-mediated lncRNA regulation, suggesting that polyadenylation or a subset of proteins bound to the poly(A) tail signals for RNA exosome-mediated lncRNA turnover (39). Furthermore, PABPN1 and PAP hyperadenylate improperly-spliced RNAs and target them for degradation by the RNA exosome (40). Other studies of Pab2 in *S. pombe* have implicated PABPN1 in RNA exosome-mediated poly(A) tail trimming during small nucleolar RNA (snoRNA) processing (43). Therefore, PABPN1, through its participation in polyadenylation, acts as a quality control mechanism to regulate noncoding RNAs and improperly-spliced RNAs by targeting these RNAs for decay.

Given that PABPN1 plays multiple critical roles in RNA processing, why *PABPN1* mutations affect only a subset of skeletal muscles is not at all clear. This muscle-specificity suggests that there are muscle-specific requirements for PABPN1 which are not fulfilled by expression of the mutant protein. Thus, studying the muscle-specific roles for PABPN1 could serve as a paradigm for understanding tissue-specific functions and requirements for ubiquitously expressed RNA binding proteins. Overall, studying RNA binding proteins in the context of tissues affected by mutations in genes encoding these RNA binding proteins could give insight into the complex tissue-specific requirements for RNA processing.

1.4 PABPN1 and OPMD: Understanding tissue-specific disease pathology

Many potential mechanisms could contribute to tissue-specific susceptibility when RNA binding proteins are perturbed. For example, RNA binding proteins could have tissue-specific roles that when disrupted cause pathology. Mutations could disrupt regulation of tissue-specific transcripts or protein-protein interactions critical for proper tissue function. Expansion mutations could cause proteins to form insoluble aggregates that sequester critical proteins and RNAs, or are themselves toxic, although this mechanism would most likely cause pathology in tissues with high levels of mutant protein. Finally, genes encoding RNA binding proteins can be dynamically regulated in a tissue-specific manner, and tissues with relatively high or low protein levels may be more susceptible to pathology than tissues with moderate levels. Overall, the mechanisms underlying tissue-specificity are likely complex and require precise investigation of

tissue-specific functions of these proteins under wildtype and disease conditions in the relevant tissue.

1.4.1 Aggregates and OPMD

Both loss and gain of function models of OPMD pathology have been proposed (33). However, little work has investigated the molecular mechanisms dictating why muscle is specifically affected by PABPN1 expansion. One pathological hallmark of OPMD is the presence of insoluble nuclear PABPN1 aggregates in patient muscle nuclei (33). Although the role aggregates play in OPMD pathogenesis is unclear, previous studies have shown that aggregate-containing nuclei from a mouse model overexpressing expanded PABPN1 express markers of cell death (56), suggesting that aggregates could be toxic. However, *in vitro* studies determined that both wildtype and expanded PABPN1 protein form aggregates *in vitro*, and that the N-terminus of the PABPN1 protein, which contains the alanine tract, is not required for aggregate formation (57). Finally, PABPN1 aggregates are observed in ~5-10% of patient myonuclei and also occur in the brain, a tissue not typically affected in OPMD. Thus, the exact mechanisms by which expanded PABPN1 aggregates contribute to OPMD pathology are unknown, and raise questions as to whether aggregate-mediated toxicity drives pathology.

Furthermore, diseases with aggregate-mediated pathology would logically occur in tissues with high levels of mutant protein. Therefore, if toxic PABPN1 aggregates were the sole mechanism of OPMD pathology, pathology would likely occur in tissues with high levels of toxic PABPN1 protein. However, experiments performed in human

and mouse tissues reveal that *PABPN1* mRNA and protein levels are low in muscle tissue relative to non-muscle tissue (58); PABPN1 levels are lower still in muscles typically affected in OPMD (58). These low PABPN1 protein levels could predispose skeletal muscle to OPMD pathology and suggest a loss of function model of OPMD pathology. Furthermore, *PABPN1* mRNA levels decrease with age in human muscle (59) so a small loss of function may become progressively worse as PABPN1 levels decline.

Because PABPN1 protein levels are so low in muscle, any reduction in functional PABPN1 could be more deleterious in this tissue compared to tissues with higher PABPN1 levels. These low levels bring forth a model where levels of functional PABPN1 fall below a threshold for OPMD pathology only in skeletal muscle while non-muscle tissues are able to fulfill critical requirements for PABPN1 function and thus do not present pathology (Figure 1.4). Multiple lines of evidence support this loss of function model of OPMD pathology. For example, aggregates from OPMD patient tissue contain PABPN1, depleting the functional pool of PABPN1. Other RNA binding proteins such as TDP-43 (60) and hnRNPA1 have also been detected in PABPN1 aggregates (61). Still other studies have demonstrated that nuclear PABPN1 aggregates can sequester polyadenylated RNAs leading to nuclear retention and RNA splicing defects (62). These data suggest that sequestration of PABPN1 and other proteins in addition to RNAs could interfere with proper localization and function of these factors. Therefore, low PABPN1 levels and hindered functionality due to this sequestration in aggregates could interfere with proper cellular function, contributing to OPMD pathology.

In vivo knockdown studies have begun to address loss versus gain of function models of OPMD pathogenesis. Both overexpression of expanded PABPN1 and *in vivo*

PABPN1 knockdown experiments (63) found similar changes in alternative polyadenylation for mRNAs encoding components of the proteasome complex, demonstrating that PABPN1 reduction results in similar 3'UTR changes as expression of alanine-expanded PABPN1.

Furthermore, overexpression of wild type PABPN1 in an expanded PABPN1 overexpression mouse model is anti-apoptotic and increases muscle strength while aggregate numbers remain unchanged (64). This result suggests that expanded PABPN1 is not able to fulfill anti-apoptotic functions independent of aggregate numbers. However, expanded PABPN1 cannot rescue splicing defects observed with PABPN1 depletion (52), suggesting that the molecular mechanisms involved in OPMD are complex.

1.4.2 Exploring loss of function models of OPMD pathology

Many of these studies utilized transgenic overexpression ~10-30 fold of expanded PABPN1 to investigate molecular events that could contribute to OPMD pathology (65). However, overexpression may alter aggregate dynamics and cause phenotypes that are not physiologically relevant as PABPN1 protein levels are low in muscle. Thus, it is important to study phenotypes caused by endogenous levels of expanded PABPN1 to identify the molecular deficiencies in muscle that could contribute to OPMD pathology.

The ideal system in which to study endogenous expression of expanded PABPN1 are muscle biopsies or primary tissue culture from OPMD patients. However, because OPMD is rare and muscle biopsies are invasive and painful, patient tissue samples are limited in availability. To begin to investigate whether endogenous levels of expanded

PABPN1 cause molecular and pathological phenotypes *in vivo*, we have generated a novel mouse model that represents the closest-available genocopy of OPMD patients. This mouse model contains one conditional expanded *Pabpn1* allele knocked-in at the native *Pabpn1* locus and one wildtype *Pabpn1* allele. This knocked-in allele encodes PABPN1 with a 17-alanine tract at the N-terminus of the protein (*Pabpn1*$^{A17/+}$). Chapter 2 details our histiological and molecular characterization of this novel OPMD mouse model, including early-onset polyadenylation defects and mitochondrial abnormalities observed in muscles from *Pabpn1*$^{A17/+}$ mice relative to muscles from wildtype animals (*Pabpn1*$^{+/+}$).

To explore if these defects are due to loss of PABPN1 function, we generated an additional mouse model with one *Pabpn1* allele deleted (*Pabpn1*$^{+/\Delta}$). By comparing phenotypes in *Pabpn1*$^{+/\Delta}$ mice with those observed in *Pabpn1*$^{A17/+}$ mice, we assessed the functionality of expanded PABPN1. Work described in Chapter 2 has identified multiple lines of evidence supporting that a partial loss of PABPN1 function contributes to OPMD pathology as similar histologic, proteomic, and molecular phenotypes are observed in muscles from both *Pabpn1*$^{A17/+}$ and *Pabpn1*$^{+/\Delta}$ mice.

1.4.3 Muscle-specific *Pabpn1* regulation: Implications for OPMD therapies

The overlapping phenotypes between *Pabpn1*$^{A17/+}$ and *Pabpn1*$^{+/\Delta}$ mice suggest that a loss of PABPN1 function could contribute to molecular deficiencies and OPMD pathology. This loss of function model of pathology suggests that increasing PABPN1 protein levels in muscle could be protective against OPMD pathology. Therefore,

understanding the mechanisms contributing to low PABPN1 protein levels specifically in muscle is critical and could identify therapeutic targets for modulating PABPN1 levels in OPMD.

Previous work has determined that the *Pabpn1* transcript is unstable in muscle relative to non-muscle tissue, contributing to low steady-state *Pabpn1* transcript and protein levels *in vivo* (58). Because the bulk of PABPN1 regulation occurs at the level of RNA, we investigated the mechanisms regulating *Pabpn1* transcript and protein levels specifically in muscle. In Chapter 3, we define multiple *cis*-regulatory elements present in the *Pabpn1* 3'UTR including an AU-rich element (ARE) that binds the RNA binding protein HuR *in vitro.* We have determined that HuR negatively regulates *Pabpn1* mRNA and protein levels in cultured muscle cells, and this regulation is conserved *in vivo.* Furthermore, Chapter 4 describes three microRNAs, miR-141-5p, miR-331-3p, and miR-532-3p, which regulate endogenous *Pabpn1* transcript levels in muscle cells. miR141-5p and miR-331-3p negatively regulate *Pabpn1* transcript levels, whereas miR-532-3p positively regulates *Pabpn1* transcript levels. We have determined that regulation by these microRNAs is dependent on the *Pabpn1* 3'UTR, suggesting that these microRNAs directly regulate *Pabpn1* transcript levels. Overall, these data demonstrate that the complex interplay between *cis*-regulatory elements and *trans*-acting factors regulate *Pabpn1* transcript levels in muscle, and these mechanisms could be exploited to develop therapies to increase PABPN1 protein levels in OPMD.

1.5 Figures

Figure 1.1: RNA processing summary

mRNAs undergo multiple steps of RNA processing. Co-transcriptional processing steps include 5' capping, splicing, and 3'end formation which is followed by polyadenyation. Improperly processed transcripts are degraded in the nucleus, while properly processed transcripts are exported through the nuclear pore complex (NPC). Once in the cytoplasm, transcripts can be translated to generate proteins or are degraded by cytoplasmic degradation machinery.

Figure 1.2: The C-terminal domain of the large subunit of RNA polymerase II

Co-transcriptional RNA processing events are linked changes in phosphorylation status of the C-terminal domain (CTD) of the large subunit of RNA polymerase II. The CTD contains heptad repeats, here illustrated as YSPTSPS. Serine 2 and Serine 5 phosphorylation status correlates with transcriptional stages: Serine 5 is phosphorylated (Ser5P) during transcriptional initiation, Serine 5 and Serine 2 are both phosphorylated (Ser2,5P) as a transitional phase during transcriptional elongation, and transcriptional termination is marked by Ser2P.

Figure 1.3: PABPN1 structure and function

A. Schematic of PABPN1 domain structure. The amino terminal domain contains a 10-alanine residue tract (top) directly following the initial methionine. This alanine tract is expanded to 11-18 alanines in OPMD (bottom). The coiled-coil domain (CCD)is required

for stimulation of poly(A) polymerase (PAP). The RNA recognition motif (RRM) mediates RNA binding. The RRM is also required for oligomerization of PABPN1 as is the C-terminal domain, which also contains a nuclear localization signal. **B.** Schematic of nuclear and cytoplasmic PABPN1 functions in a metazoan cell. The functions of PABPN1 include stimulating PAP, regulating poly(A) tail length by affecting the interaction between PAP and the cleavage and polyadenylation specificity factor (CPSF), and influencing poly (A) signal (PAS) selection. PABPN1 affects exosome-mediated degradation of non-coding RNAs as well as poly(A) RNA export. Other nuclear functions include regulation of splicing and *PABPN1* autoregulation through intron retention and degradation by affecting binding of splicing factors SRSF10 and HNRNPH/A2. Cytoplasmic functions include participation in the pioneer round of translation via interaction with cap binding proteins CBP20 and CBP80 as well as deadenylation and decay via the CCR4-NOT RNA degradation machinery.

Figure 1.4: A threshold model of OPMD pathology

Under non-disease conditions (N, black bars), nuclear poly(A) binding protein (PABPN1) levels are higher in non-muscle tissues than muscle tissues. If expansion mutations in the *PABPN1* gene in oculopharyngeal muscular dystrophy (OPMD, white bars) reduce PABPN1 protein function, function could be reduced in all tissues, with levels falling below a critical threshold for pathology only in muscle, resulting in pathology specifically in this tissue.

Figure 1.1

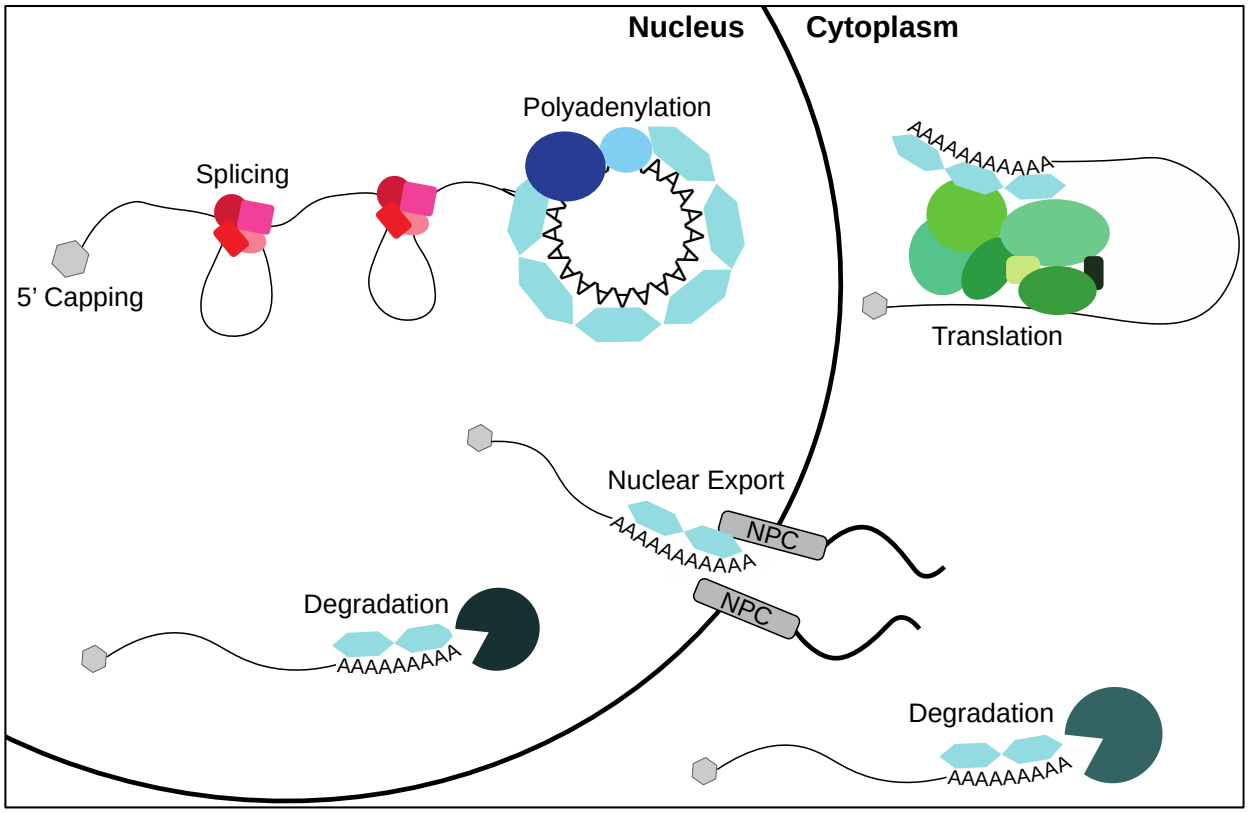

Nucleus
Cytoplasm
Polyadenylation
Splicing
5' Capping
AAAAAAAAAA
Translation
Nuclear Export
AAAAAAAAAA
NPC
NPC
Degradation
AAAAAAAAA
Degradation
AAAAAAAAA

Figure 1.2

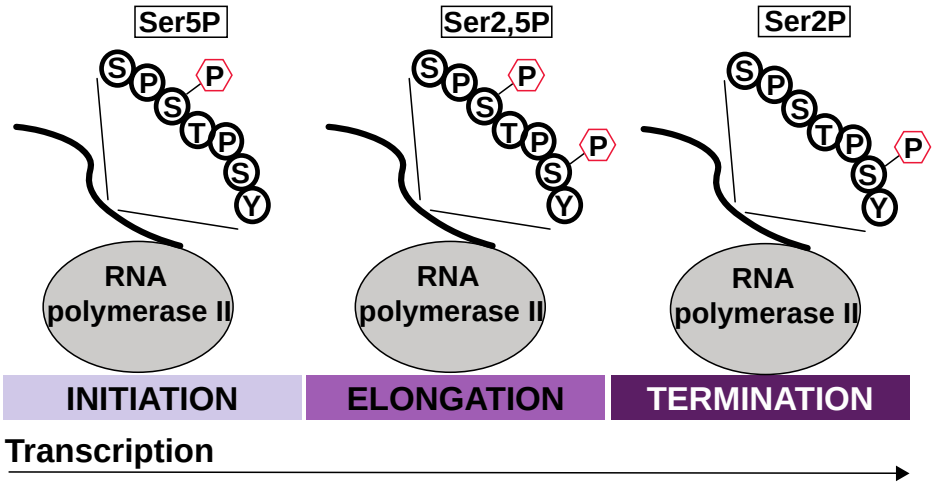

Ser5P
Ser2,5P
Ser2P
S P S T P S Y
P
RNA polymerase II
INITIATION
ELONGATION
TERMINATION
Transcription

Figure 1.3

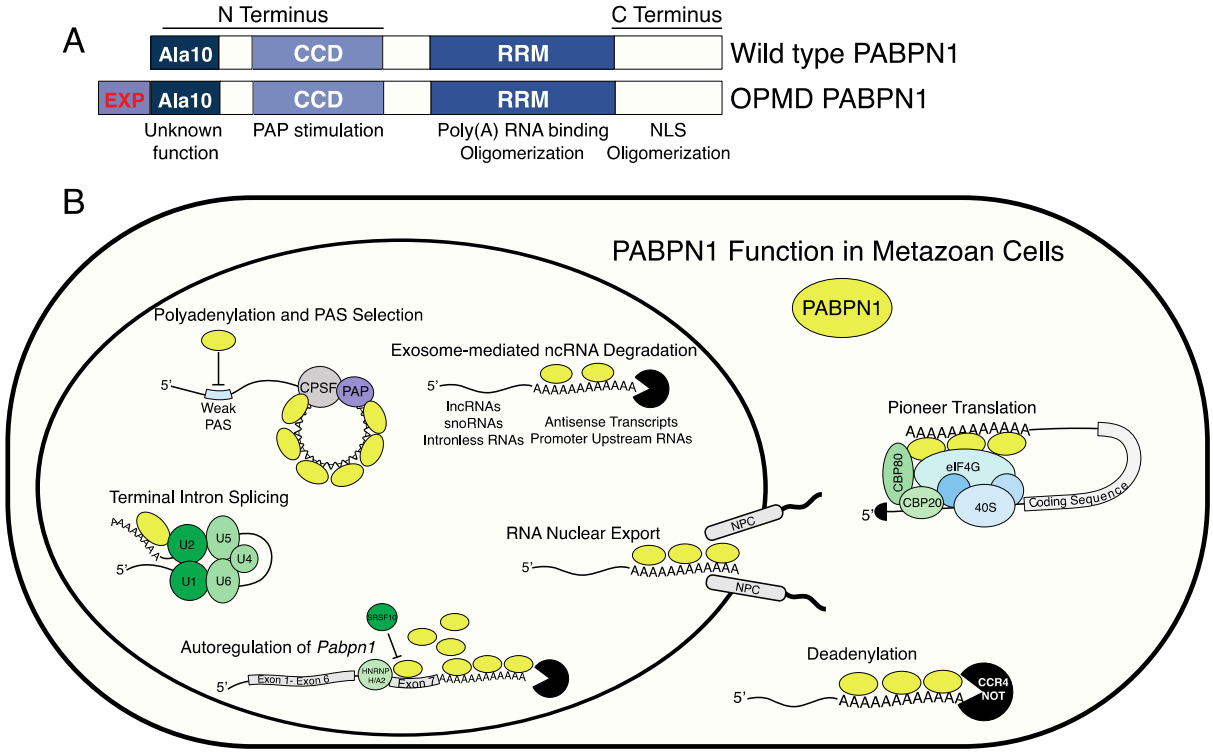
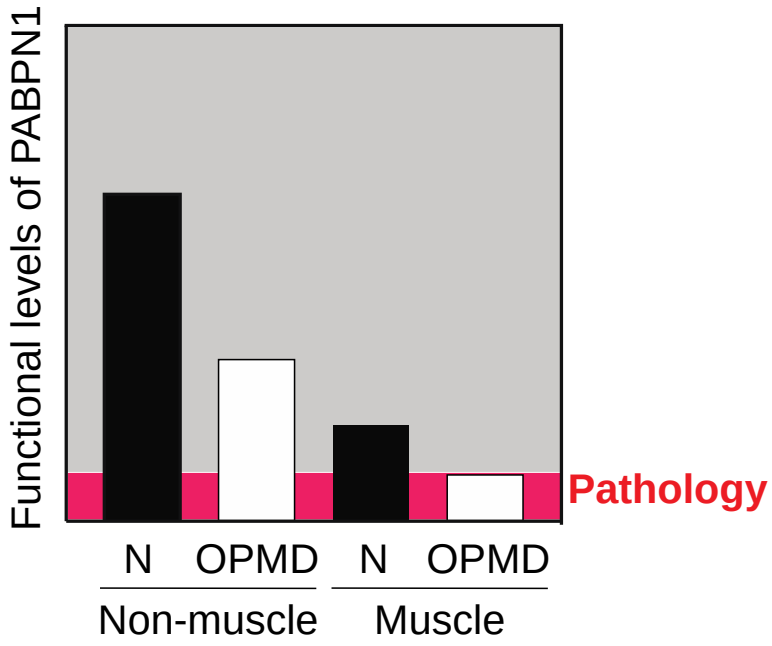

A
N Terminus
C Terminus
Ala10
CCD
RRM
Wild type PABPN1
EXP
Ala10
CCD
RRM
OPMD PABPN1
Unknown function
PAP stimulation
Poly(A) RNA binding Oligomerization
NLS Oligomerization
B
PABPN1 Function in Metazoan Cells
PABPN1
Polyadenylation and PAS Selection
Weak PAS
CPSF
PAP
Exosome-mediated ncRNA Degradation
IncRNAs
snoRNAs
Intronless RNAs
Antisense Transcripts
Promoter Upstream RNAs
Pioneer Translation
CBP80
eIF4G
CBP20
40S
Coding Sequence
Terminal Intron Splicing
U2
U5
U4
U1
U6
RNA Nuclear Export
NPC
NPC
Autoregulation of Pabpn1
hnRNP H/A2
Exon 1- Exon 6
Exon 7
Deadenylation
CCR4 NOT

Figure 1.4


Functional levels of PABPN1
Pathology
N
OPMD
N
OPMD
Non-muscle
Muscle

Chapter 2: Novel mouse models of oculopharyngeal muscular dystrophy (OPMD) reveal early onset mitochondrial defects and suggest loss of PABPN1 may contribute to pathology

This chapter is adapted from the following published work:

Vest K.E., **Phillips B.L.**, Banerjee A., Apponi L.H., Dammer E.B., Xu W., Zheng D., Yu J., Tian B., Pavlath G.K., Corbett A.H. *Novel mouse models of oculopharyngeal muscular dystrophy (OPMD) reveal early onset mitochondrial defects and suggest loss of PABPN1 may contribute to pathology.* Human Molecular Genetics (2017) **26**(17):3235-3252. Published by Oxford University Press.

Brittany Phillips performed experiments and data analyses for the following figures: Figure 2.6, 2.8, Supplementary Figure 2.6, Supplementary Figure 2.7, and Supplementary Figure 2.8.

2.1 Summary

Oculopharyngeal muscular dystrophy (OPMD) is a late onset disease caused by polyalanine expansion in the poly(A) binding protein nuclear 1 (PABPN1). Several mouse models have been generated to study OPMD; however, most of these models have employed transgenic overexpression of alanine-expanded PABPN1. These models do not recapitulate the OPMD patient genotype and PABPN1 overexpression could confound molecular phenotypes. We have developed a knock-in mouse model of OPMD (*Pabpn1*$^{+/A17}$) that contains one alanine-expanded *Pabpn1* allele under the control of the native promoter and one wild-type *Pabpn1* allele. This mouse is the closest available genocopy of OPMD patients. We show that *Pabpn1*$^{+/A17}$ mice have a mild myopathic phenotype in adult and aged animals. We examined early molecular and biochemical phenotypes associated with expressing native levels of A17-PABPN1 and detected shorter poly(A) tails, modest changes in poly(A) signal (PAS) usage, and evidence of mitochondrial damage in these mice. Recent studies have suggested that a loss of PABPN1 function could contribute to muscle pathology in OPMD. To investigate a loss of function model of pathology, we generated a heterozygous *Pabpn1* knock-out mouse model (*Pabpn1*$^{+/\Delta}$). Like the *Pabpn1*$^{+/A17}$ mice, *Pabpn1*$^{+/\Delta}$ mice have mild histologic defects, shorter poly(A) tails, and evidence of mitochondrial damage. However, the phenotypes detected in *Pabpn1*$^{+/\Delta}$ mice only partially overlap with those detected in *Pabpn1*$^{+/A17}$ mice. These results suggest that loss of PABPN1 function could contribute to but may not completely explain the pathology detected in *Pabpn1*$^{+/A17}$ mice.

2.2 Introduction

Oculopharyngeal muscular dystrophy (OPMD) is a late onset disease that causes ptosis, dysphagia, and loss of mobility due to progressive weakness of eyelid, pharyngeal, and proximal limb muscles. OPMD is not fatal but associated symptoms greatly affect quality of life. Dysphagia is often the prognostic symptom and the associated loss of nutrition and recurring aspiration pneumonia can be fatal (66,67). Treatments for OPMD involve surgical intervention to raise eyelids and loosen constrictor muscles of the throat to ease swallowing. No pharmacologic interventions are currently available.

The vast majority of OPMD patients have a dominant *GCG* expansion in the gene encoding the polyadenosine binding protein nuclear 1 (PABPN1) (38). Unlike other disease-causing expansion mutations, the dominant OPMD mutations are rather modest. Wild type PABPN1 contains a stretch of ten alanine residues that occurs just after the initial methionine. This alanine tract, which occurs in a domain with unknown function, is expanded to 11-18 alanine residues in patients with OPMD (38,68,69). PABPN1 is a ubiquitously expressed RNA binding protein that is most well characterized as a regulator of polyadenosine (poly[A]) tail length and alternative polyadenylation (APA) as well as expression of antisense RNAs associated with gene promoters (41,42,44,70-72). Why a small expansion of alanine residues in a ubiquitously expressed RNA binding protein causes a late-onset muscular dystrophy affecting specific muscle groups is not known.

The pathological hallmark identified in OPMD patient tissues is the presence of nuclear aggregates (73) that contain PABPN1 and other proteins including additional RNA binding proteins and heat shock proteins (61,74,75). OPMD has been considered an aggregate disease akin to other triplet repeat diseases such as Huntington's disease. If

toxic aggregates were the sole cause of pathology in OPMD, tissues that express the highest levels of the aggregate-prone protein would likely be most susceptible to pathology. However, muscle contains very low levels of PABPN1 relative to unaffected tissues (58), suggesting a more complex scenario that cannot be solely explained by aggregate-mediated toxicity. Consistent with a more complex mechanism underlying pathology, nuclear aggregates are found only in a small subset of myonuclei in OPMD patients (73,76,77) and have been detected in unaffected muscles and brain from an OPMD patient (78). Dynamic, non-pathogenic PABPN1 nuclear aggregates have also been detected in rat neurons (79,80), demonstrating that wild type PABPN1 can form aggregates. However, several studies link the presence of PABPN1 aggregates to cell death (56,64,81,82), supporting a role for aggregates contributing to pathology. Indeed, aggregation may deplete the functional PABPN1 pool. This model is supported by the fact that overexpression of wild type PABPN1 reverses apoptotic cell death caused by overexpression of alanine-expanded PABPN1 despite an increase in the number of nuclear aggregates (64).

Challenges exist for studying the functional consequences of *PABPN1* mutation in OPMD patient muscle tissues. Patient tissues are difficult to obtain as OPMD is a rare muscle disease, affecting approximately 1:100,000 individuals (83), and muscle biopsy is invasive and painful. Furthermore, the late disease onset makes it difficult to perform systematic and controlled molecular studies with multiple time points and muscles in OPMD patient samples. Given the limited availability of patient samples, several model systems have been developed in which to study OPMD pathology including cell culture, *Caenorhabditis elegans*, *Drosophila melanogaster*, and mouse (56,75,78,84-89).

However, these models have typically employed tissue-specific and/or constitutive promoters that lead to very high levels of expanded PABPN1 that do not account for the low expression of PABPN1 in muscle and thus complicate interpretation. Therefore, it is important to analyze the effects of endogenous levels of expanded PABPN1 in order understand how altered alanine-expanded PABPN1 causes muscle-specific pathology.

Many studies of OPMD pathology have been performed in mouse models. The most commonly used mouse model contains a transgenic expanded bovine *PABPN1* allele under the control of the muscle-specific human skeletal actin promoter. These animals express 17 alanine-expanded PABPN1 at levels 10 to 30-fold over endogenous wild type PABPN1 levels (56,65). In these mice, approximately 5% of myonuclei contain intranuclear aggregates at three months of age and about 20% of myonuclei contain aggregates at 6 months of age. The high percentage of aggregate-positive nuclei and subsequent apoptotic myofiber death make this mouse an excellent model in which to study drug treatments that ameliorate aggregate formation and apoptosis. Indeed, several promising drugs have been tested in this model and results from some of these molecular and biochemical studies have been confirmed in OPMD patient samples (59,90-92).

Despite the utility of previous mouse models (56,81,82) in drug discovery, this model does not recapitulate the patient genotype and because PABPN1 has many important roles in RNA processing, overexpression of A17-PABPN1 may confound molecular data obtained from these studies. For example, *in vitro* experiments showed that wild type PABPN1 protein can form aggregates (93) and additional studies revealed that muscle-specific overexpression of wild type PABPN1 causes a functional phenotype in flies (87). While patient tissues would be the ideal system in which to define the early

molecular events associated with OPMD pathology, they are limited in availability. Thus, a model system that closely mimics the genotype of OPMD patients is needed.

We have developed a novel mouse model of OPMD that represents the closest available genocopy of OPMD patients. This new OPMD model contains a conditional *Pabpn1* knock-in allele (*Pabpn1*^{A17}) at the native *Pabpn1* locus, which encodes a 17 alanine-expanded PABPN1 protein (A17 PABPN1). Histologic analysis revealed that *Pabpn1*$^{+/A17}$ mice have age-related and muscle-specific defects. We performed molecular and biochemical studies in young *Pabpn1*$^{+/A17}$ mice and detected early-onset defects in polyadenylation and mitochondrial abnormalities. We also compared *Pabpn1*$^{+/A17}$ mice to a heterozygous *Pabpn1* knockout mouse and found similar but not identical histologic, proteomic, and molecular phenotypes. Taken together these data suggest that loss of PABPN1 function could contribute to the muscle pathology detected in *Pabpn1*$^{+/A17}$ mice.

2.3 Results

2.3.1 Generation of a new mouse model of OPMD

To generate the closest possible model to a patient genotype in autosomal dominant OPMD, we created a mouse model with a single *(GCG)*$_7$ expanded *Pabpn1* allele knocked in at the native *Pabpn1* locus. As shown in Figure 2.1A, the targeted allele contains a cassette with murine *Pabpn1* cDNA, an internal ribosome entry site (IRES), *eGFP* and a neomycin resistance gene (*Neo*) flanked by loxP sites followed by a 3' arm containing the (*GCG*)$_7$ expansion. Following flippase and Cre-mediated recombination, the portion of the targeted allele between loxP sites is excised and the knock-in allele that

encodes a 17-alanine expanded PABPN1 protein (A17 PABPN1) remains at the native *Pabpn1* locus. To create mice that express A17 PABPN1 throughout development, mice containing the targeted allele were crossed with mice expressing Cre recombinase using the EIIa promoter (EIIA Cre mice), which is activated during early embryogenesis (94). Thus, resulting progeny (*Pabpn1*$^{+/A17}$) contain a single wild type *Pabpn1* allele and a knocked-in expanded *Pabpn1*^{A17} allele in all tissues, including germline tissues. These mice with the recombined allele were then crossed with each other to generate experimental animals that did not contain Cre recombinase. Mouse genotypes were determined by PCR on tail snips at weaning and in muscle tissue after tissue harvest (Figure S1A,B).

To confirm that these mice express A17 PABPN1 in muscle and non-muscle tissue, we analyzed A17 PABPN1 expression by immunoblotting using an antibody specific to the alanine expansion (95) (Figure 2.1B). As steady-state PABPN1 protein levels are low in muscle (50), we analyzed primary myoblasts and spleen tissue, which both have higher levels of PABPN1 than muscle tissue. We detected alanine-expanded PABPN1 in both myoblasts and spleen collected from *Pabpn1*$^{+/A17}$ mice but not *Pabpn1*$^{+/+}$ mice. As a control, we included samples from an established transgenic OPMD mouse model (56), which shows high A17-PABPN1 levels as well as total PABPN1 levels consistent with the overexpression of transgenic alanine-expanded PABPN1 in this model. To assess whether total PABPN1 levels differ in the *Pabpn1*$^{+/A17}$ mice compared to the *Pabpn1*$^{+/+}$ mice, we analyzed total PABPN1 protein in muscle and spleen (Figure 2.1C, D). This analysis revealed that, in contrast to the transgenic mouse model (56), PABPN1 protein levels are similar in *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice in

both muscle and non-muscle tissue. We also detected comparable levels of *Pabpn1* mRNA in muscle and non-muscle tissue from *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice when analyzed by quantitative reverse transcriptase PCR (qRT-PCR) (Figure S1C). KCl-insoluble PABPN1 aggregates are a hallmark of OPMD pathology that are found in muscle sections from OPMD patients and animal models of OPMD (33). To determine whether muscle from the *Pabpn1*$^{+/A17}$ mice show this hallmark of OPMD, we used immunofluorescence staining to detect PABPN1 (Figure 2.1E) in KCl-insoluble aggregates in rectus femoris (RF) muscles from, *Pabpn1*$^{+/A17}$ mice but not in *Pabpn1*$^{+/+}$ mice. To demonstrate that KCl-insoluble aggregates are detected in multiple muscles, we also use immunofluorescence staining to detect aggregates in tibialis anterior (TA) muscle (not shown). Quantification of immunostaining results from TA muscles revealed that, in *Pabpn1*$^{+/A17}$ mice, ~7% of myonuclei contain KCl-insoluble PABPN1 aggregates at three months of age and ~ 10% of myonuclei contain KCl-insoluble aggregates at 18 months of age (Figure 2.1F), though the difference between the two ages was not statistically significant. Taken together, these results demonstrate that we have generated a new mouse model of OPMD.

To assess functional phenotypes in the *Pabpn1*$^{+/A17}$ mice, we employed standard assays including rotorod latency to fall and grip strength. Separate cohorts of male and female mice were tested at 6, 12, and 18 months of age. The *Pabpn1*$^{+/A17}$ mice did not have any significant defects as determined by these measurements (data not shown). However, the *Pabpn1*$^{+/A17}$ mice did display hind limb clasping starting at nine months of age (Figure S2). This phenotype is consistent with a previously reported transgenic mouse model of OPMD that contains an expanded human *PABPN1* gene under the

control of the *PABPN1* native promoter (78). To produce a more severe phenotype, we attempted to generate *Pabpn1*$^{A17/A17}$ mice by crossing two *Pabpn1*$^{+/A17}$ mice. We used PCR primers designed to amplify differently sized products for the *Pabpn1*$^{+}$ and *Pabpn1*^{A17} alleles to test for the presence of *Pabpn1*$^{A17/A17}$ mice (Figure S3A). Analysis of genotyping results revealed that only one of >40 pups tested positive for this genotype at weaning (Figure S3B). This one mouse likely tested positive due to some tissue mosaicism, possibly due to uneven Cre expression during embryonic development, as the animal had no overt phenotype. We found that litters generated from these crosses were significantly smaller by ~20% than crosses of *Pabpn1*$^{+/A17}$ mice with *Pabpn1*$^{+/+}$ mice, suggesting that *Pabpn1*$^{A17/A17}$ mice may be rarely generated or do not survive to weaning age (Figure S3C).

2.3.2 *Pabpn1*$^{+/A17}$ knock-in mice have smaller muscles than wild type littermates

To evaluate muscle histopathology, we cross-sectioned representative proximal, rectus femoris (RF), and distal, tibialis anterior (TA), muscles from male *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/+}$ mice and stained with hematoxylin and eosin (H&E). Results from histologic analyses are summarized in Figure 2.2 (RF) and Figure S4 (TA). To take into consideration the late onset of OPMD, we analyzed muscles from animals at 6 and 18 months of age. No gross histologic defects were detected in *Pabpn1*$^{+/A17}$ mice aside from some small myofibers present in RF (Figure 2.2A,D) and TA muscles (Figure S4 A,D). However, cross-sectional area (CSA) of *Pabpn1*$^{+/A17}$ myofibers was significantly smaller in both muscles at 6 months of age (Figure 2.2B, Figure S4B). In 18-month-old *Pabpn1*$^{+/A17}$ mice, RF myofiber CSA (Figure 2.2E) was smaller, but TA myofiber CSA was larger (Figure S4E). We quantified the percentage of central-nucleated fibers (CNF),

a marker of myofiber regeneration (96), and detected a significant increase in CNF in TA muscles at 6 months of age in the *Pabpn1*$^{+/A17}$ mice (Figure S4C) but no change in TA muscles at 18 months of age (Figure S4F) or in RF muscles (Figure 2.2C,F) at any age. These results demonstrate small but significant defects in muscles of *Pabpn1*$^{+/A17}$ mice compared to *Pabpn1*$^{+/+}$ mice that change with age in a muscle-specific manner.

2.3.3 *Pabpn1*$^{+/A17}$ mice have subtle early-onset RNA phenotypes

PABPN1 has multiple roles in RNA processing, but the most well-described function is enhancing the processivity of poly(A) polymerase to modulate the length of poly(A) tails added to RNAs (70-72). To determine whether polyadenylation is altered in *Pabpn1*$^{+/A17}$ mice prior to pathology, we measured bulk poly(A) tail length of total RNA isolated from RF muscles of three-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. We used an established assay where 3' end radio labeled poly(A) tracts are resolved by polyacrylamide gel electrophoresis (33,50,97). As shown in Figure 2.3A, bulk poly(A) tail length is decreased in RNA from *Pabpn1*$^{+/A17}$ mice as compared to *Pabpn1*$^{+/+}$ mice. Densitometry analysis at regions representing short, medium, and long poly(A) tails revealed that the longest poly(A) tails were decreased by ~20% in *Pabpn1*$^{+/A17}$ mice relative to *Pabpn1*$^{+/+}$ mice (Figure 2.3B). Similarly, densitometry scans of each lane of poly(A) tails revealed a significant decrease in overall poly(A) tail length in *Pabpn1*$^{+/A17}$ mice (Figure 2.3C). Taken together, these results indicate global poly(A) tail shortening in *Pabpn1*$^{+/A17}$ mice. This result demonstrates for the first time that expression of native levels of A17 PABPN affects polyadenylation, one of the major functions of PABPN1.

Another important function of PABPN1 in RNA processing is regulation of alternate polyadenylation signal (PAS) utilization, known as alternative polyadenylation

(APA). APA can affect the coding region of the transcript or the length of the 3' untranslated region (UTR) depending on the cleavage site employed. Previous studies demonstrated that knockdown of *PABPN1* in C2C12, HeLa and U2OS cell lines or transgenic overexpression of alanine-expanded PABPN1 in mouse muscles cause a global shift in PAS usage from distal to proximal PAS sites (41,42,44). To determine whether expression of native levels of A17 PABPN1 affects global PAS usage, we performed 3' region extraction and deep sequencing (3' READS) (98) on total RNA isolated from RF muscles of three-month-old male *Pabpn1*$^{+/A17}$ mice. As shown in Figure 2.4A, scatter plots representing the log ratio of distal or proximal PAS reads revealed no global shift toward either distal or proximal PAS utilization. However, a modest change in PAS usage (toward either distal or proximal PAS) in a small number of transcripts was detected (Figure 2.4A), suggesting that native levels of A17 PABPN1 do affect PAS utilization. Heat maps of PAS reads (Figure 2.4B) revealed that samples grouped by genotype, suggesting that changes detected in PAS usage were related to A17 PABPN1 expression rather than being stochastic or caused by some unknown variable. However, these data also indicate that the change in APA is very modest. The complete list of genes that show a change in PAS utilization compared to control can be found in Table S2.

To validate the 3’ READS data, we selected a subset of candidates to test for altered PAS usage using qRT-PCR primers targeting the coding sequence to amplify total transcript for each candidate and primers targeting the region between the proximal and distal PAS to amplify transcripts that were polyadenylated at the distal PAS (Figure 2.4C, top). APA was detected as a shift in the ratio of distal PAS transcripts to the total transcript pool. We detected altered PAS usage in three candidate transcripts, *Ddit4l*,

Slc25a4, and *Mcts1* in *Pabpn1*$^{+/A17}$ mice as demonstrated by the altered ratio of transcripts utilizing the distal PAS to the total pool of transcripts (Figure 2.4C, bottom). Two candidates had a corresponding change in steady-state levels when assayed by qRT-PCR and normalized to samples from *Pabpn1*$^{+/+}$ mice (Figure 2.4D). Thus, expression of native levels of A17 PABPN1 has a modest effect on PABPN1 function in alternative polyadenylation, shortening and lengthening 3' UTRs of a subset of mRNAs as compared to *Pabpn1*$^{+/+}$ mice.

The 3' READS data also provide information about steady-state mRNA levels. As summarized in the Venn diagram in Figure 2.4E, of 4640 gene-level reads detected in *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/+}$ mice, 95 were depleted from *Pabpn1*$^{+/A17}$ mice and 30 were enriched in *Pabpn1*$^{+/A17}$ mice. A heat map of gene-level reads demonstrates that *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/+}$ mice group by genotype (Figure 2.4F). We performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (99,100). This analysis revealed that genes with decreased steady-state mRNA levels in *Pabpn1*$^{+/A17}$ mice are enriched for terms involved in mitochondrial metabolism (Figure 2.4G) suggesting there could be altered mitochondrial function in these mice.

2.3.4 Proteomic analysis reveals mitochondrial defects in *Pabpn1*$^{+/A17}$ mice

To determine whether results obtained from the 3' READS experiment were reflected in the proteome of *Pabpn1*$^{+/A17}$ mice, we analyzed global proteomic changes comparing RF muscle from three-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice by mass spectrometry. Label-free quantification of proteins was performed by MaxQuant analysis of extracted ion currents (XIC). The average XIC across three biological repeats

was calculated for *Pabpn1*$^{+/A17}$ mice and their corresponding *Pabpn1*$^{+/+}$ littermates. To determine which proteins were altered, we calculated the log2 ratio of XIC values from *Pabpn1*$^{+/A17}$ relative to *Pabpn1*$^{+/+}$ mice. As represented in Figure 2.5A, of the 2423 proteins detected in all samples, 115 proteins were less abundant in *Pabpn1*$^{+/A17}$ mice and 147 proteins were more abundant in *Pabpn1*$^{+/A17}$ mice relative to the corresponding *Pabpn1*$^{+/+}$ mice. A summary of the total proteomic data for *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice is presented in Table S3. We performed GO and KEGG pathway analysis using DAVID. The results revealed that proteins that are depleted in *Pabpn1*$^{+/A17}$ relative to *Pabpn1*$^{+/+}$ are most significantly enriched for terms involving mitochondria and mitochondrial metabolism (Figure 2.5B) while proteins that were more abundant in *Pabpn1*$^{+/A17}$ relative to *Pabpn1*$^{+/+}$ were most significantly enriched for terms involving cell structure and the proteasome complex (Figure 2.5C).

Many of the mitochondrial proteins depleted in *Pabpn1*$^{+/A17}$ mice are involved in the electron transport chain and oxidative phosphorylation (OXPHOS). To test for dysregulation of OXPHOS proteins in *Pabpn1*$^{+/A17}$ mice, we performed immunoblot analysis for electron transport chain complex subunits (Figure 2.5D). While each complex of the electron transport chain contains multiple subunits, many of the subunits are labile when not incorporated into the complex. Thus, levels of labile subunits can be used as a proxy to quantify the entire complex. We detected significant decreases in SDHB (Complex II), COX1 (Complex IV), and ATP5A (Complex V) proteins in RF muscles from three-month-old male *Pabpn1*$^{+/A17}$ mice relative to *Pabpn1*$^{+/+}$ mice. Trends toward decreased levels of NDUFB8 (Complex I) and UQCR2 (Complex III) were also detected. To determine whether the decrease in OXPHOS proteins was due to a decrease

in total mitochondrial content, we quantified total mitochondria by immunoblotting for the outer membrane voltage-dependent anion channel (VDAC) and by performing quantitative PCR (qPCR) to detect mitochondrial DNA (mtDNA). No change was detected in total mitochondrial content as assessed by immunoblotting for VDAC (Figure 2.5E). However, Figure 2.5F shows that a trend toward increased levels of two mitochondrial genes, *mtCox1* and *mtCox3* (101), was detected when analyzed by qPCR and normalized to two nuclear genes, *Hprt* and *Dio3*, suggesting increased levels of mtDNA relative to nuclear DNA. Variability in samples from *Pabpn1*$^{+/A17}$ mice, potentially due to varying degrees of mitochondrial damage, likely contributed to the lack of statistical significance for these changes.

The fact that increased mtDNA level does not correspond with an increased mitochondrial protein marker level suggests an increase in mtDNA replication, which can indicate mitochondrial dysfunction related to age (102). To further test for aberrant mitochondrial function in *Pabpn1*$^{+/A17}$ mice we tested for defects in electron transport chain complex activity. We stained sectioned RF and TA muscles from three-month-old male mice for succinate dehydrogenase activity (SDH, Complex II) using nitro blue tetrazolium assay. We detected a significant decrease in SDH positive fibers in RF muscles from *Pabpn1*$^{+/A17}$ mice (Figure S5) but did not detect any change in SDH positive fibers in TA muscles from *Pabpn1*$^{+/A17}$ mice. This result suggests that functional mitochondrial defects are muscle-specific and are more prevalent in proximal limb muscle than distal limb muscle. To further assess mitochondrial function, we performed *in vitro* assays of NADH:ubiquinone oxidoreductase (Complex I) activity using homogenize RF muscle from 3, 12, and 18-month-old *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice.

We detected a trend toward decreased Complex I activity in RF muscle homogenate from 3-month-old male mice and a statistically significant decrease in 12-month-old male *Pabpn1*$^{+/A17}$ mice (Figure S5C). These results suggest that early-onset mitochondrial defects are exacerbated with age. These results support the proteomic and immunoblot data and indicate functional mitochondrial defects in *Pabpn1*$^{+/A17}$ mice.

2.3.5 Comparison with *Pabpn1* knock-out suggests loss of function may contribute to but does not fully explain pathology

Recent studies have suggested that a loss of PABPN1 function could contribute to phenotypes associated with OPMD (50,58,63-65). To determine whether a loss of PABPN1 function contributes to the phenotypes detected in *Pabpn1*$^{+/A17}$ mice, we generated *Pabpn1* knock-out mice for comparison. As shown in Figure 2.6A, the *Pabpn1* targeted allele contains regions homologous to exons 1 and 2 of the *Pabpn1* gene and a puromycin resistance gene within two FRT sites flanked by *loxP* sites. Following flippase and Cre-mediated recombination, exons 1 and 2 are excised resulting in a null *Pabpn1* allele with no transcriptional start site. To generate mice with *Pabpn1* knocked out from early embryogenesis, mice containing the targeted allele were crossed with EIIA Cre mice to generate *Pabpn1*$^{+/\Delta}$ mice. Mouse genotype was determined by PCR of tail snips taken at weaning and by PCR of muscle samples after tissue harvest (Figure S6A,B). When *Pabpn1*$^{+/\Delta}$ mice were crossed with *Pabpn1*$^{+/\Delta}$ mice to generate homozygous *Pabpn1* knock-out mice, none were detected at weaning (Figure S6C). Furthermore, quantification of total pups produced by crossing *Pabpn1*$^{+/\Delta}$ mice to *Pabpn1*$^{+/\Delta}$ mice revealed that litters produced by these crosses were significantly smaller than those

generated by crossing *Pabpn1*$^{+/\Delta}$ mice to *Pabpn1*$^{+/+}$ (Figure S6D). These data show that *Pabpn1*$^{\Delta/\Delta}$ mice may not be viable and suggest that *Pabpn1* may be an essential gene. All subsequent studies to explore the functional consequences decreased PABPN1 were carried out using *Pabpn1*$^{+/\Delta}$ mice. Analysis of qRT-PCR and immunoblot data confirmed that muscle and non-muscle tissue from *Pabpn1*$^{+/\Delta}$ mice show a ~50% decrease in steady-state *Pabpn1* mRNA and PABPN1 protein when compared to *Pabpn1*$^{+/+}$ littermates (Figure 2.6B,C), validating this as a model to assess the consequence of decreased levels of PABPN1.

To test for muscle histopathology associated with decreased PABPN1 levels, we analyzed H&E stained sections of RF and TA muscles from male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice at 6 and 18 months of age. In six-month-old *Pabpn1*$^{+/\Delta}$ mice, we detected some small myofibers in both RF and TA muscles but no other obvious defects (Figure 2.6D,G and Figure S7A,D). Frequency distribution of myofiber CSA of both RF and TA from six-month-old *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice (Figure 2.6E and Figure S7B) revealed significant shifts toward smaller myofibers in *Pabpn1*$^{+/\Delta}$ mice in both muscles. These results are consistent with defects detected in six-month-old *Pabpn1*$^{+/A17}$ mice (Figures 2.2 and S4). In 18-month-old *Pabpn1*$^{+/\Delta}$ mice, CSA was significantly increased in both RF (Figure 2.6G,H) and TA (Figure S7D,E) myofibers indicating age-related hypertrophy. In contrast, this age-related hypertrophy was detected only in TA muscles from 18-month-old *Pabpn1*$^{+/A17}$ mice. CNF were not significantly increased in RF or TA muscles (Figure 2.6F, I Figure S7C,F) of *Pabpn1*$^{+/\Delta}$ mice relative to *Pabpn1*$^{+/+}$ mice at 6 months of age whereas CNF were increased in TA muscles of *Pabpn1*$^{+/A17}$ mice at this age (Figure S4C). These results indicate defects in *Pabpn1*$^{+/\Delta}$ muscles are similar to those

detected in *Pabpn1*$^{+/A17}$ muscles in adult animals. However, aged *Pabpn1*$^{+/\Delta}$ mice do not recapitulate all histologic defects detected in aged *Pabpn1*$^{+/A17}$ mice.

To determine whether *Pabpn1*$^{+/\Delta}$ mice had similar molecular phenotypes to *Pabpn1*$^{+/A17}$ mice, we analyzed RF muscles isolated from three-month-old male *Pabpn1*$^{+/\Delta}$ and *Pabpn1*$^{+/+}$ mice. Quantification of poly(A) tail length revealed a ~ 50% decrease in the longest poly(A) tails in the *Pabpn1*$^{+/\Delta}$ mice (Figures 2.7A,B), which is comparable to the 50% decrease in longest poly(A) tails caused by *Pabpn1* knockdown in primary myoblasts (50). Densitometry analysis of each lane indicates a global decrease in bulk poly(A) tail length in RNA isolated from RF muscles of *Pabpn1*$^{+/\Delta}$ mice (Figure 2.7C). This result is similar to the global poly(A) tail shortening detected in primary myoblasts after Pabpn1 knockdown (50), and to the poly(A) tail shortening detected in *Pabpn1*$^{+/A17}$ mice (Figure 2.3). To determine whether loss of 50% of PABPN1 affects APA, we performed 3' READS analysis comparing *Pabpn1*$^{+/\Delta}$ and *Pabpn1*$^{+/+}$ mice. Scatter plots of distal to proximal PAS reads shown in Figure 2.7C revealed very little change in PAS usage with 50% PABPN1 depletion. Cluster analysis of PAS reads revealed no grouping of *Pabpn1*$^{+/\Delta}$ and *Pabpn1*$^{+/+}$ samples by genotype (Figure 2.7D), indicating no significant change in PAS usage on a global scale. These data indicate that PABPN1 depletion and expression of native levels of A17 PABPN1 both lead to shorter poly(A) tails but 50% depletion of PABPN1 does not cause a global shift in PAS utilization at least in young mice.

To explore whether the proteomic changes we identified in *Pabpn1*$^{+/A17}$ mice are also detected in *Pabpn1*$^{+/\Delta}$ mice, we used global proteomic analysis of RF muscles comparing three-month-old *Pabpn1*$^{+/\Delta}$ to *Pabpn1*$^{+/+}$ mice (Figure 2.8A). We detected a

total of 1147 proteins with 94 proteins depleted from and 165 proteins enriched in *Pabpn1*$^{+/\Delta}$ mice relative to *Pabpn1*$^{+/+}$. Analysis of GO terms and KEGG pathways revealed that several terms related to mitochondrial metabolism and the proteasome complex were altered in RF muscles from *Pabpn1*$^{+/\Delta}$ mice relative to *Pabpn1*$^{+/+}$ mice (Figure 2.8B,C), which was similar to changes detected in *Pabpn1*$^{+/A17}$ mice (compare to Figure 2.5B,C). To determine whether mitochondrial dysregulation occurs in RF muscles from *Pabpn1*$^{+/\Delta}$ mice, we examined levels of electron transport chain complex subunits by immunoblotting and examining total mitochondrial content in *Pabpn1*$^{+/\Delta}$ and *Pabpn1*$^{+/+}$ mice. Representative immunoblot and densitometric quantification revealed that, in contrast to *Pabpn1*$^{+/A17}$ mice, RF muscles from *Pabpn1*$^{+/\Delta}$ had significantly increased steady-state levels of SDHB while levels of UQCR2 and ATP5A trended upward but this increase was not statistically significant (Figure 2.8D,E). Analysis of SDH activity revealed that no significant change was detected in the number of SDH positive myofibers in TA or RF (Figure S8). However, consistent with data showing increased levels of electron transport chain complex proteins (Figure 2.8D,E), a trend toward an increase in SDH positive myofibers was detected in RF muscles. These results contrast with the data showing decreased SDH positive myofibers in RF muscles of *Pabpn1*$^{+/A17}$ mice (Figure S5). No significant change in total mitochondrial content was detected by immunoblot analysis of VDAC levels (Figure 2.8F), though levels trended downward. Analysis of qPCR data revealed decreased levels of *mtCox1* and *mtCox3* relative to nuclear DNA markers (Figure 2.8G), indicating decreased mtDNA content whereas *Pabpn1*$^{+/A17}$ mice showed increased mtDNA content (compare to Figure 2.6C).

These data suggest that different biochemical consequences result from similar molecular phenotypes detected in *Pabpn1*$^{+/\Delta}$ mice and *Pabpn1*$^{+/A17}$ mice.

2.4 Discussion

We have generated a mouse model that is the closest available genocopy to OPMD patients. This mouse model stands out in comparison to other models in that the expanded *Pabpn1* allele is present at the native locus and thus the model does not rely on transgenic expression of alanine-expanded PABPN1 and does not overexpress alanine-expanded PABPN1. In *Pabpn1*$^{+/A17}$ mice we detected small but significant muscle-specific and age-related muscle histopathology. We also demonstrated that expression of native levels of A17-PABPN1 causes an overall decrease in poly(A) tail length. Although we did not detect a global shift in PAS, we detected modest changes in APA in a subset of RNAs. We also demonstrated that mitochondrial defects occur at the RNA and protein level. Comparing *Pabpn1*$^{+/\Delta}$ and *Pabpn1*$^{+/A17}$ mice revealed a partial overlap in histologic and molecular phenotypes, supporting a model where a partial loss of PABPN1 function contributes to but does not fully explain muscle pathology.

These novel mouse models that allow analysis of mammalian PABPN1 function *in vivo*. We were unable to generate homozygous *Pabpn1*$^{\Delta/\Delta}$ mice, which suggests that *Pabpn1* may be an essential gene in mice. We also found that we were unable to generate homozygous *Pabpn1*$^{A17/A17}$ mice in Mendelian ratios, suggesting that A17-PABPN1 may not be able to perform essential PABPN1 functions. However, because aggregate-mediated PABPN1 sequestration or cell death may be exacerbated by increased A17-PABPN1 doses, it is unclear whether decreased PABPN1 function, PABPN1 aggregation, or some combination of aggregation and loss of function underlie this

lethality. Considering that Ala13 homozygous OPMD patients have been reported (103), perhaps a mouse model containing *Pabpn1*$^{A13/A13}$ would be a useful system in which to study the effects of expanded PABPN1 dosage.

To assess muscle histopathology in these newly generated mouse models, we performed histologic analysis on both proximal (RF) and distal (TA) limb muscles from six-month-old (adult) and 18-month-old (aged) mice. We detected atrophy in RF muscles of adult and aged *Pabpn1*$^{+/A17}$ mice along with small and central nucleated fibers (CNF), though no significant change in total CNF was detected. These results are consistent with proximal limb involvement reported in OPMD patients (104,105). We also detected atrophy in TA muscles of adult *Pabpn1*$^{+/A17}$ mice but detected hypertrophy in TA muscles of aged *Pabpn1*$^{+/A17}$ mice. Age-related hypertrophy may be a result of a compensatory myogenic response, which is supported by the finding of increased CNF in TA muscles of adult *Pabpn1*$^{+/A17}$ mice. Age-related hypertrophy has also been detected in other mouse models of muscle disease such as the *mdx* model of Duchenne muscular dystrophy (106,107). Atrophy was also detected in RF and TA muscles from adult *Pabpn1*$^{+/\Delta}$ mice. However, in aged *Pabpn1*$^{+/\Delta}$ mice, hypertrophy was detected in both RF and TA muscles. These data indicate phenotypic overlap in adult *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice that diverges with age in proximal (RF) but not distal (TA) muscles.

To examine molecular events that could precede pathology in *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice, we analyzed RF muscle isolated from three-month-old animals. Decreased bulk poly(A) tail length detected in muscles from three-month-old *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice is an intriguing result that suggests that alanine-expanded PABPN1 is not fully functional in this canonical role for PABPN1. Previous work has shown that

Pabpn1 knockdown in cell culture results in a similar decrease in poly(A) tail length (50), so this result could be due depletion of the total pool of PABPN1 in *Pabpn1*$^{+/\Delta}$ mice. Our analysis shows that steady state levels of PABPN1 protein are comparable in muscles from *Pabpn1*$^{+/A17}$ mice and *Pabpn1*$^{+/+}$ mice, suggesting that the poly(A) tail defect is caused by loss of PABPN1 function rather than PABPN1 protein. However, the decrease in bulk poly(A) tail length in the *Pabpn1*$^{+/A17}$ mice could result from sequestration of PABPN1 protein in aggregates with no impact on total protein levels detected. This possibility of loss of PABPN1 function due to sequestration complicates studies to assess the function of alanine-expanded PABPN1.

We did not detect a global shift in alternative polyadenylation signal (PAS) usage (APA) in muscles from *Pabpn1*$^{+/A17}$ mice. Previous reports linking PABPN1 to global changes in APA with a shift toward more proximal PAS usage were performed in cell culture or animal models that either had *Pabpn1* knocked down or transgenic alanine-expanded PABPN1 overexpressed (41,42,44). Though these studies point to an important function of PABPN1 in regulating PAS selection, our results suggest that this function is only modestly disrupted by expressing native levels of alanine-expanded PABPN1 in young *Pabpn1*$^{+/A17}$ mice. Also notable is no discernable regulation of antisense transcripts from gene promoters in *Pabpn1*$^{+/A17}$ mice, which was reported previously with *Pabpn1* knockdown in cell culture (44). Therefore, aberrant PAS utilization on a global scale may not be an early driver of pathology associated with OPMD. Although small changes in PAS utilization were detected in *Pabpn1*$^{+/A17}$ mice, even fewer were detected in *Pabpn1*$^{+/\Delta}$ mice. A previous study of mouse TA muscle with viral-delivered shRNA-mediated acute *Pabpn1* knockdown demonstrated altered PAS utilization in transcripts

corresponding to a subset of proteasome genes (63) so perhaps in the case of some RNAs, 50% loss of PABPN1 is sufficient to alter PAS usage. However, in this model, while the overall knockdown achieved was ~50%, comparable to our genetic model, the viral delivery could have led to heterogeneity in the amount of PABPN1 knockdown achieved in different muscle fibers. Because it has been shown in humans that steady-state *PABPN1* mRNA levels decrease with age (59), perhaps the effect of alanine-expanded PABPN1 on PAS utilization becomes more relevant to OPMD pathology over time. Further work is needed to determine whether changes in PAS utilization increase with age in *Pabpn1*$^{+/A17}$ or *Pabpn1*$^{+/\Delta}$ mice.

Quantification of steady-state mRNA levels from the 3'READS experiment as well as validated proteomic data showed that mitochondria are impaired in *Pabpn1*$^{+/A17}$ mice. These results are consistent with those reported previously in a study implicating mitochondrial dysfunction in *Drosophila* and mouse models of OPMD that overexpress alanine-expanded PABPN1 as well as in OPMD patient tissues (90). In those studies, decreased poly(A) tail length of key nuclear-encoded mitochondrial mRNAs was also reported. We detected altered PAS usage and increased levels of *Ddit4l* mRNA, which is a regulator of mTOR signaling that may contribute to impaired mitochondrial function. Elevated levels of *Ddit4l* have been reported in cases of disuse atrophy (108,109) so decreased myofiber CSA detected in RF of adult and aged mice could be a result of increased *Ddit4l* and decreased mTOR signaling. We detected altered levels of nuclear-encoded mitochondrial mRNAs and protein but did not detect any change in levels of VDAC protein, a marker of total mitochondrial content. We did detect increased mitochondrial genome copy number in *Pabpn1*$^{+/A17}$ mice, which is associated with aging

in skeletal muscle (102). Increased mitochondrial DNA detected in *Pabpn1*$^{+/A17}$ mice suggests that premature mitochondrial aging may be an early-onset symptom of OPMD.

Interestingly, we detected different effects on OXPHOS protein levels and mtDNA content in *Pabpn1*$^{+/A17}$ mice and *Pabpn1*$^{+/\Delta}$ mice. Previous reports of mitochondrial defects have been attributed to shortened poly(A) tails in key nuclear-encoded mitochondrial RNAs (90). However, the fact that both *Pabpn1*$^{+/A17}$ mice and *Pabpn1*$^{+/\Delta}$ mice had shortened bulk poly(A) tails but different mitochondrial phenotypes suggests that another mechanism may contribute to mitochondrial defects. A recent report demonstrated that transgenic expression of alanine-expanded PABPN1 causes accumulation of misspliced *TNNT3* mRNA due to *TNNT3* mRNA sequestration in aggregates away from SC35-containing splicing speckles in samples from a cell culture model of OPMD as well as samples from OPMD patients (62). Perhaps other aberrant splicing events in nuclear-encoded mitochondrial genes could occur in *Pabpn1*$^{+/A17}$ mice and contribute to aberrant mitochondrial function.

While patient tissues are the ideal system in which to study mechanisms of OPMD pathology, they present a challenge in that their limited availability, combined with the late onset nature of the disease, makes controlled studies of symptomatic and pre-symptomatic individuals difficult. This mouse model, which is the closest available genocopy to OPMD patients, can be used to perform the types of studies that are hindered by limited patient tissue availability. Critically, comparison of molecular and biochemical data generated in *Pabpn1*$^{+/A17}$ mice recapitulate some of the major results obtained from OPMD patient tissue studies. For example, the validated 3' READS and proteomics data supporting mitochondrial dysfunction in young *Pabpn1*$^{+/A17}$ mice agree with proteomic

data reported from OPMD patient biopsies showing mitochondrial proteins as a major category of dysregulated proteins (90). Other microarray studies have shown altered levels of mRNAs encoding mitochondrial proteins in OPMD patient tissues (59) and defects in mitochondrial morphology have been reported in several patient case studies (110-113). With the availability of these mice, future studies can determine the mechanistic details of mitochondrial defects.

In addition to our study, several additional reports suggest that loss of PABPN1 function could contribute to muscle pathology in OPMD (50,63-65). If alanine expansion in PABPN1 interferes with normal PABPN1 function(s), overexpression of the wildtype PABPN1 should ameliorate deleterious effects. Indeed, overexpression of wildtype PABPN1 is protective against depletion of the X-linked inhibitor of apoptosis and cell death in mice overexpressing A17-PABPN1 (64) and recent studies of potential gene therapies for OPMD revealed that expression of exogenous wild type *Pabpn1* reversed the myofiber cross-sectional area defect in muscles from the A17.1 mouse (114). Furthermore, over-expression of wildtype PABPN1 prevents functional defects in pharyngeal muscles associated with aging in mice (65). In fact, loss of protein function may be more relevant than previously appreciated in other aggregate-associated diseases. For example, loss of APP/Aβ protein function may contribute to Alzheimer's disease pathology (115) and loss of huntingtin protein during development is associated with age-related neurodegeneration in a mouse model (116). Future work is needed to elucidate the complicated interplay between toxicity, cell death, and loss of protein function in aggregate-associated proteinopathies. These studies would benefit from use of model

systems such as the one described here that do not employ transgenic overexpression of pathogenic proteins.

OPMD is often characterized as a disease of aggregate-mediated toxicity, though new data are emerging to argue that this model is too simplistic to explain the pathology that occurs in a subset of muscles in OPMD patients. Several studies performed in a well-established mouse model that overexpresses alanine-expanded PABPN1 demonstrate that aggregate formation leads to severe muscle phenotypes and that removal of aggregates reverses those phenotypes (56,81,82). However, a different study reported that soluble expanded PABPN1 was more toxic to cells than the aggregated form (117) and still another reported that overexpression of wild type PABPN1 ameliorated phenotypes related to alanine-expanded PABPN1 overexpression (64). Data from our newly-created *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice indicate some molecular and histologic overlap but also divergence in biochemical phenotypes. The fact that phenotypes in *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice do not overlap completely hint at a more complex model of pathology in which defects due to loss of function and PABPN1 aggregation are not mutually exclusive. With the creation of this mouse model, a new tool is available for future studies such as determining tissue-specific requirements for PABPN1, assessing the function of alanine-expanded PABPN1 in muscle, and exploring whether native levels of alanine-expanded PABPN1 form aggregates in other tissues. These studies will help to determine how alanine-expanded PABPN1 expression causes pathology in skeletal muscle and may help to identify new targets for the treatment of OPMD.

2.5 Materials and Methods

Generation of *Pabpn1*$^{+/A17}$ knock-in mice (Figure 2.1A): The conditional *Pabpn1*$^{+/A17}$ knock-in mouse line was generated by Ozgene Pty Ltd (Bentley WA, Australia). A targeting construct was cloned to contain the $(GCG)_7$ expansion at the 5' end. Upstream of the expansion mutation, the targeting construct contained a wild type murine *Pabpn1* cDNA sequence, an internal ribosome entry site (*IRES*), a gene encoding enhanced green fluorescent protein (*eGFP*) and a neomycin resistance gene (*Neo*) flanked by flippase recognition target (*FRT*) sites. The wild type *Pabpn1* cDNA, *IRES*, *eGFP*, and *Neo* were flanked by *loxP* sites. The entire targeting construct was flanked at each end with regions of homology to the 5' end of the *Pabpn1* coding sequence and the region immediately upstream of the *Pabpn1* coding sequence. Targeting constructs were electroporated into a C57BL/6 ES cell line, Bruce4 (118). Homologous recombinant ES cell clones were identified by Southern hybridization and injected into blastocysts. Male chimeric mice were obtained and crossed to C57BL/6J females to establish heterozygous offspring. Mice heterozygous for the targeted allele were crossed with mice expressing flippase to remove the *Neo* gene and subsequently crossed with mice containing the Cre recombinase gene under the control of the *EIIa* promoter (The Jackson Laboratory, Bar Harbor, ME) to generate mice containing *Pabpn1*$^{+/A17}$ in all cells including the germline (Figure 2.1A) (94). *Pabpn1*$^{+/A17}$ knock-in mice were genotyped by PCR using primers designed to detect the presence of the *loxP* site (A17 Allele primers) and the presence of the *eGFP* gene (Floxed Allele primers) found in the floxed allele (Figure S1). Animals that generated a PCR product for the *loxP* site but not for the *eGFP* gene were considered

positive for the *Pabpn1*^{A17} allele. Initial crosses were tested for the presence of the Cre recombinase gene using *Cre* primers designed by the Jackson Laboratory (oIMR1084 and oIMR1085). To determine whether knock-in mice were homozygous or heterozygous for *Pabpn1*^{A17}, primers were designed to amplify differently sized products for the *Pabpn1*$^{+}$ and *Pabpn1*^{A17} alleles (WT/A17 primers, Figure S3A). Animals that generated products for both *Pabpn1*$^{+}$ and *Pabpn1*^{A17} alleles were considered to be heterozygous knock-ins. Sequences for PCR primers used for genotyping can be found in Table S1. Experiments were performed using male mice at 3, 6, 9, or 18 months of age. All animal experiments were performed in accordance with approved guidelines and ethical approval from Emory University's Institutional Animal Care and Use Committee and in compliance with the National Institutes of Health.

Generation of conditional *Pabpn1* knock-out mouse (Figure 2.3A): Generation of the *Pabpn1* conditional knockout allele was performed using standard methods (119). The targeting construct contained regions of homology to exons one and two of the *Pabpn1* gene, a puromycin resistance gene (*Puro*) flanked by *FRT* sequences between two *loxP* sites. Targeting constructs were cloned into the pFlexible vector (119), which was linearized by *KpnI* digest and electroporated into HZ1.1 mouse embryonic stem cells (C57BL/6 background) at the Emory University School of Medicine Transgenic Mouse and Gene Targeting Core Facility. Puromycin resistant clones were screened by Southern blot to identify clones with proper construct integration. Selected clones were expanded, injected into blastocysts, and implanted into C57Bl/6 mice. Mice heterozygous for the targeted allele were crossed with mice expressing flippase to remove the *Neo* gene and subsequently crossed with mice containing the Cre recombinase gene under the control of

the *EIIa* promoter (Jackson Laboratories) to generate mice with one *Pabpn1* allele knocked out in all cells including the germline (Figure 2.3A). Heterozygous *Pabpn1* knock-out mice were genotyped by PCR using a set of three primers (WT/Floxed Allele and Knockout Allele primers) designed to amplify differently-sized products from the wild type, floxed, and knock-out alleles. Animals that were positive for the knockout allele and wild type allele but not the floxed allele were considered heterozygous *Pabpn1* knock-out animals. Sequences for PCR primers used for genotyping can be found in Table S1. Experiments were performed using mice at 3, 6, or 18 months of age.

Real-time PCR analysis: For analysis of steady-state mRNA levels, RNA was isolated from rectus femoris and tibialis anterior muscles, or non-muscle tissue from three-month-old mice using TRIzol (ThermoFisher Scientific) according to the manufacturer's instructions. From isolated RNA, cDNA was synthesized using the M-MLV reverse transcriptase kit (Invitrogen) and Rnasin plus (Promega). For analysis of mitochondrial and nuclear DNA markers, total DNA was isolated from muscles using phenol chloroform: isoamyl alcohol (25:24:1, ThermoFisher Scientific) according to the manufacturer's instructions. For quantitative PCR analysis, approximately 10 ng of cDNA or genomic DNA samples were mixed with appropriate primers (Table S1) and SYBR Select Master Mix (Applied Biosciences). All samples were analyzed using the comparative Ct method (120) on an Applied Biosciences Step One Plus Real Time PCR System. For analysis of alternative polyadenylation, primers were designed to amplify a region within the coding sequence of the mRNA (Total) or the region between the distal and proximal PAS (Distal). PAS usage was quantified by calculating ΔΔCt as Distal-Total and fold change was reported as a ratio of $Pabpn1^{+/+}$ controls. For analysis of

cDNA, all samples were normalized to *18S rRNA* or *Gapdh*. For analysis of mtDNA, samples were normalized to two nuclear reference genes, *Hprt* and *Dio3*.

Immunoblotting: From three-month-old mice, rectus femoris and tibialis anterior muscles were either homogenized or primary mouse cells were isolated and cultured, and lysed by standard methods (50). Samples were separated by SDS-PAGE on Mini-PROTEAN TGX Any kD polyacrylamide gels (BioRad), and transferred to 0.2 μm nitrocellulose membrane (BioRad). Membranes were blocked in either 5% bovine serum albumin (BSA) in Tris-buffered saline pH 7.4, with 0.1% Tween-20 (TBS-T) to be probed with α-VDAC antibodies or with 5% nonfat milk in TBS-T for all other immunoblots. The following primary antibodies were used: α-PABPN1 (1:8000) (50), α-Alanine expansion (1:2000) (95), α-HSP90 (1:5000, Santa Cruz Biotechnology #sc-13119), α-OXPHOS antibody cocktail (1:4000, Abcam #ab110413), and α-VDAC (1:4000, Abcam #15895). In all cases, following primary antibody incubation, immunoblots were incubated in the appropriate horseradish peroxidase conjugated secondary α-IgG antibodies (1:4000, Jackson ImmunoResearch Laboratories #115-036-062, #711-036-152).

Preparation of muscle tissue for histologic analysis: Tibialis anterior and rectus femoris muscles from six or eighteen-month-old mice were dissected and flash frozen in cryo-freezing medium (Sakura OCT-TissueTek). Muscles were cut into 14 μm serial cross-sections using a cryostat (Leica Biosystems) and sections were mounted on Superfrost Plus slides (Fisher Scientific). Sectioned muscles were either immediately stained with hematoxylin and eosin (H&E) or frozen for later analysis.

Succinate dehydrogenase activity assay: Sectioned tibialis anterior and rectus femoris muscles from three-month-old mice were stained for succinate dehydrogenase activity as described (121). Briefly, muscle sections were incubated in staining solution (1.5 mM nitroblue tetrazolium, 130 mM sodium succinate, 0.2 mM phenazine methosulfate, 1 mM sodium azide in PBS) at 37 ˚C for 20 minutes, washed in PBS, dehydrated in ethanol and xylene, and mounted with Permount mounting medium (Fisher Scientific).

Complex I Activity Assay: Activity assay kits (Cayman Chemical) were used to quantify mitochondrial electron transport chain Complex I activity in homogenized rectus femoris muscle according to the manufacturer's instructions. Briefly, whole RF muscles isolated from 3, 12, and 18-month-old male mice were homogenized on ice in muscle homogenization buffer (122) using a Dounce homogenizer and insoluble debris was removed by centrifugation. Complex I activity was measured as oxidation of NADH by measuring a decrease in Absorbance at 340 nm using a BioTek Neo plate reader.

Histologic analyses: For analysis of H&E and SDH stained slides, stained sections were imaged in the same anatomic region across all genotypes and ages. Exact localization within the muscle was determined using blood vessels and nerves as landmarks. For analysis of myofiber cross-sectional area (CSA), images were obtained from four separate sections representing the proximal, middle, and distal sections of each muscle for each biological replicate. CSA was measured by manually outlining individual myofibers using ImageJ (NIH). For each biological replicate, the CSA of 400-500 myofibers was quantified. For quantification of central nucleated fibers (CNF), the number of myofibers with one or more centrally-located nuclei was quantified and reported as a percentage of total myofibers analyzed. For SDH analysis, images were

obtained for 6-8 separate sections representing the proximal, middle, and distal portion of the muscle. SDH positive myofibers were counted using ImageJ and calculated as a percentage of total myofibers for 500-700 myofibers analyzed per biological replicate.

Immunostaining: For detection of PABPN1, rectus femoris or tibialis anterior muscle sections from three-month-old mice were permeabilized in 0.3% Triton x-100 in PBS at room temperature (15 minutes), incubated in blocking buffer (0.3% BSA, 0.1% Triton x-100 in PBS) at room temperature (1-2 hours) and incubated in primary α-PABPN1 diluted 1:1000 in 0.5x blocking buffer in PBS at 4 ˚C (overnight). Slides were washed in 0.5x blocking buffer and incubated in FITC-conjugated donkey anti-rabbit IgG (Jackson ImmunoResearch Laboratories #711-096-152) diluted 1:500 in 0.5x blocking buffer at room temperature (1 hour). Slides were washed in 0.5x blocking buffer and incubated in 4'6-diamidino-2-phenylindole (DAPI, 1 μg/ml in PBS) to visualize nuclei, and mounted with Vectashield anti-fade mounting medium (Vector Laboratories). To remove soluble proteins, slides were incubated at room temperature in 1 M KCl in HPEM buffer (30 mM HEPES, 65 mM PIPES, 10 mM EDTA, 2 mM $MgCl_2$, pH 6.9) for one hour after permeabilization and before blocking.

Quantification of PABPN1 aggregates: KCl-insoluble aggregates were quantified as percentage of aggregates positive for PABPN1 by immunostaining in KCl treated sections relative to total DAPI-stained nuclei. For each genotype and age, ≥ 12 sections for n = 3 animals were quantified.

Determination of bulk poly(A) tail length: Analysis of poly(A) tails from bulk RNA was performed as described (50). Briefly, 10 μg of RNA isolated from rectus femoris muscles of three-month-old male mice was labeled with [^{32}P]-cytidine 3',5' bisphosphate

using T4 RNA ligase and non-poly(A) RNA was digested using RNase A and T1. Labeled RNA was separated on a 7% polyacrylamide gel containing 8 M urea, 90 mM tris-borate, and 2 mM EDTA and visualized by autoradiography. Poly(A) tail length was quantified by densitometric analysis of gels using ImageJ. Different areas on images of poly(A) tail gels were chosen to represent short, medium, and long poly(A) tails. Integrated density was quantified and normalized to the integrated density of a region representing very short poly(A) tails that is not expected to be affected by loss of PABPN1 function. Entire lanes were also scanned and data for each genotype was binned by poly(A) tail length using GraphPad Prism 7.0. Statistical significance was determined using one-way ANOVA. Changes in bulk poly(A) tail length were detected as a decrease in the density in regions representing longer poly(A) tails. Statistical significance was determined by non-parametric t-test using GraphPad Prism 7.0.

3' region extraction and deep sequencing (3' READS):

3' READS analysis was performed on total RNA isolated from rectus femoris muscles of three-month-old mice as described (98,123). Briefly, 1 μg of input RNA was used for each sample, and poly(A)+ RNA was selected using oligo d(T)25 magnetic beads (NEB), followed by on-bead fragmentation using RNase III (NEB). Poly(A)+ RNA fragments were then selected using a chimeric oligo containing 15 regular dTs and five locked nucleic acid dTs conjugated on streptavidin beads, followed by RNase H (NEB) digestion. RNA fragments with ligated with 5’ and 3’ adapters were reverse-transcribed, followed by PCR (15x) to obtain cDNA libraries for sequencing on the Illumina platform. Processing of 3’READS data was carried out as previously described (98). Briefly, reads were mapped to the mouse genome using bowtie 2 (124). Reads with ≥ 2 unaligned Ts at

the 5' end were used to identify PASs. PASs located within 24 nt from each other were clustered together.

APA analysis. Differential expression of APA isoforms were carried out with Fisher's exact test. Significant events were those with $p < 0.05$ and delta relative abundance >5%. Relative expression of two most abundant 3'UTR APA isoforms, e.g., proximal and distal PASs, was calculated by log2 (Distal PAS/Proximal PAS).

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis of Whole Muscle Tissue: Rectus femoris muscles were dissected from three-month-old mice and analyzed by LC-MS/MS as described (125). Muscles were homogenized in urea lysis buffer (8M urea, 100 mM $NaHPO_4$) with Pierce 1x HALT protease and phosphatase inhibitor cocktail (ThermoFisher) on a Bullet Blender (Next Advance) at 4 ˚C according to the manufacturer's instructions. Samples were then sonicated (Sonic Dismembrator, Fisher Scientific) and protein concentration was measured by bicinchoninic assay (BCA) (126). Protein homogenate was diluted in 50 mM NH_4HCO and treated with 1 mM dithiothreitol (30 minutes at room temperature) and 5 mM iodoacetamide (20 minutes at room temperature protected from light). Proteins were digested for two hours in 1:100 lysyl endopeptidase at room temperature and overnight in 1:50 trypsin (Promega) at room temperature. Digested peptides were removed from buffers on a Sep-Pak C18 column (Waters) and dried under vacuum. Dried peptides (2 µg) were resuspended in loading buffer (0.1% formic acid, 0.03% trifluoroacetic acid, 1% acetonitrile) with 0.2 pmol of isotope-labeled peptide calibrants (Life Technologies #88321). Resuspended peptides were separated on a C18 column (25 cm x 75 µm internal diameter, New Objective) by a NanoAcquity UPLC (Waters) and monitored on an Orbitrap Fusion Tribrid Mass

Spectrometer with Electron Transfer Dissociation (Thermo Scientific). Separation was performed over a 120-minute gradient (400 nl/min) of 3-80% buffer B (0.1% formic acid 0.5% DMSO in acetonitrile) into buffer A (0.1% formic acid 5% DMSO in water). The mass spectrometer cycle was programmed to collect one full MS scan followed by MS/MS scans. MS scans were collected at m/z 200 at a resolution of 70,000 in profile mode and MS/MS scans were collected at 200 m/z at a resolution of 17,500. Dynamic exclusion was set to exclude previously sequenced precursor ions for 30 seconds within a 10 ppm window. Also excluded were precursor ions with +1 or ≥ +6 charge states.

Label-Free Protein Quantification (LFQ): LFQ was performed as previously described (127,128). Raw sample data were analyzed using MaxQuant v1.5.3.30 with Thermo Foundation 2.0 for RAW file reading capability. The search engine Andromeda, integrated into MaxQuant, was used to search a concatenated target-decoy Uniprot mouse reference protein database (retrieved August, 2015; 54,489 target sequences), plus 245 contaminant proteins from the common repository of adventitious proteins (cRAP) built into MaxQuant. Methionine oxidation (+15.9949 Da), asparagine and glutamine deamidation (+0.9840 Da), and protein N-terminal acetylation (+42.0106 Da) were variable modifications (up to 5 allowed per peptide); cysteine was assigned a fixed carbamidomethyl modification (+57.0215 Da). Only fully tryptic peptides were considered with up to 2 miscleavages in the database search. A precursor mass tolerance of ±20 ppm was applied prior to mass accuracy calibration and ±4.5 ppm after internal MaxQuant calibration. Other search settings included a maximum peptide mass of 6000 Da, a minimum peptide length of 6 residues, and 0.6 Da tolerance for low resolution (ITMS) MS/MS scans. Co-fragmented peptide search was enabled to deconvolute

multiplex spectra. The false discovery rate (FDR) for peptide spectral matches, and site decoy fraction were set to 1 percent. Quantification settings were as follows: requantify with a second peak finding attempt after protein identification has completed; match MS1 peaks between runs; a 0.7 min retention time match window was used after an alignment function was found with a 20 minute RT search space. Quantitation of proteins was performed using the LFQ algorithm built into MaxQuant, which normalizes for technical variations across samples. LFQ minimum ratio count allowed was 1. The quantitation method only considered razor plus unique peptides for protein level quantitation. The full list of parameters used for MaxQuant are available upon request as a parameter.XML file. To determine changes in steady-state protein levels between muscle samples, the $\log_2$ ratio of XIC from *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ samples relative to XIC of the wild type littermates was compared. Proteins with a $\log_2$ ratio of ≥ 0.5 were considered to be enriched in *Pabpn1*$^{+/A17}$ or *Pabpn1*$^{+/\Delta}$ while proteins with a $\log_2$ ratio of ≤ -0.5 were considered to be depleted in *Pabpn1*$^{+/A17}$ knock-in or *Pabpn1*$^{+/\Delta}$ knock-out samples. Gene ontology (GO) term and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment was determined using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (99,100).

Microscopy: All stained slides were viewed using an Axioplan microscope with either a 0.5 NA 10x Plan-Neoflaur or a 0.5 NA 20x Plan-Neoflaur objective lens (Carl Zeiss MicroImaging, Inc.). Digital images were obtained with a charge-couple device camera (Carl Zeiss MicroImaging, Inc.) and Scion Image 1.63 (Scion Corp., NIH). All images were assembled and equally processed for color levels, brightness, and contrast as needed using Photoshop CC 2015 (Adobe).

Statistical analysis: All statistical analyses were performed using GraphPad Prism 5 or GraphPad Prism 7. For analysis of frequency distributions of myofiber CSA, Kruskal-Wallace test plus Dunn's post-test was performed. In all other cases, statistical significance was determined by either Student's t-test or one-way analysis of variance (ANOVA) unless otherwise noted. In all cases, samples with $p < 0.05$ were considered to be statistically significant.

2.6 Figures

Legends to Figures

Figure 2.1: Generation of the *Pabpn1*$^{+/A17}$ knock-in mouse

A) Schematic of *Pabpn1* knock-in allele. The wild type *Pabpn1* locus is shown at the top. Before Cre-mediated recombination, the Targeted Allele contains murine wild type *Pabpn1* cDNA sequence, an internal ribosome entry site (*IRES*), a gene encoding enhanced green fluorescent protein (*eGFP*) and a neomycin resistance gene (*Neo*) flanked by *loxP* sites. Downstream of the second *loxP* site is a start codon followed by a $(GCG)_7$ stretch encoding the alanine expansion at the amino terminus of PABPN1. After flippase and Cre-mediated recombination, the Mutant *Pabpn1*^{A17} Allele contains the $(GCG)_7$ expansion in Exon 1 upstream of the remaining *Pabpn1* gene along with one remaining *loxP* site. *B)* A17-PABPN1 was detected as a 50 kDa band in samples from 3-month-old male *Pabpn1* mutant *(+/A17)* mice but not *Pabpn1* wild type (+/+) mice using an α-Alanine antibody in lysates prepared from myoblasts and spleen isolated from *Pabpn1*$^{+/A17}$ mice. Lysate prepared from muscles overexpressing A17-PABPN1 (56) were used as a positive control (A17.1) for the α-Alanine antibody. *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$

samples were probed with α-PABPN1 antibody and α-HSP-90 antibody as controls. *C)* Representative immunoblot using α-PABPN1 antibody in muscle and spleen lysates from 3-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice demonstrating no change in total PABPN1 levels. Samples were probed with α-HSP-90 as a control. *D)* Quantification of PABPN1 protein levels normalized to HSP-90 protein from muscle and spleen lysates from 3-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. Data are mean ± SEM of n = 3 animals per genotype. *E)* Representative immunostaining for PABPN1 in rectus femoris muscle sections from 3-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice with and without KCl incubation. Nuclei are labeled with DAPI. KCl-insoluble aggregates are marked with white arrowheads. Bar = 25 μm. *F)* Quantification of percent KCl-insoluble aggregate nuclei in tibialis anterior muscle relative to total DAPI-stained nuclei from 3, and 18-month-old male *Pabpn1*$^{+/A17}$ mice. Data are mean ± SEM of n = 3 *Pabpn1*$^{+/A17}$ animals per age.

Figure 2.2: Smaller myofibers detected in rectus femoris muscles of *Pabpn1*$^{+/A17}$ mice

A) Histologic sections from rectus femoris (RF) muscles isolated from six-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice shown at low and high magnification. Bar = 50 μm. *B)* Significantly smaller myofiber cross-sectional area (CSA) was detected in six-month-old *Pabpn1*$^{+/A17}$ mice. Data are the frequency distribution of myofiber CSA. *C)* No significant difference occurred in the percentage of centrally nucleated fibers (CNF) in six-month-old *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. Data are mean ± SEM. *D)* RF muscle sections from 18-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice shown at low and high magnification. Bar = 50 μm. *E)* Significantly smaller CSA detected in 18-month-old

Pabpn1$^{+/A17}$ mice. Data are the frequency distribution of myofiber CSA. *F)* No significant difference occurred in the percentage of CNF in 18-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. Data are mean ± SEM. In all cases, n = 3-5 mice per genotype.

Figure 2.3: Shortened poly(A) tails in *Pabpn1*$^{+/A17}$ mice

A) Poly(A) tail length was analyzed by polyacrylamide gel electrophoresis of 3' end labeled poly(A) tracts from total RNA isolated from rectus femoris muscles of three-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. n = 3 per genotype. **p<0.05 B*) Quantification of regions representing short (1), medium (2), and long (3) poly(A) tails noted in (*A*) normalized to the total short poly(A) tails marked in (*A*) (Normalize) revealed fewer long poly(A) tails in *Pabpn1*$^{+/A17}$ mice. *C*) Quantification of gel from *(A)* by plotting density along each lane revealed shorter poly(A) tails in *Pabpn1*$^{+/A17}$ mice.

Figure 2.4: 3' READS reveals modest changes in poly(A) signal usage in *Pabpn1*$^{+/A17}$ mice

A) Scatter plot depicting that few changes were detected in poly(A) sequence (PAS) selection by 3' READS performed on RNA isolated from rectus femoris (RF) muscles from three-month-old male *Pabpn1*$^{+/A17}$ mice. Data are $\log_2$ ratio of distal to proximal PAS usage with increased distal PAS represented in red and increased proximal PAS represented in blue. n = 2 mice per genotype. *B)* Sample clustering based on PAS usage as detected by 3'READS performed on RNA isolated from RF muscles of three-month-old mice showing that *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice group by genotype. *C)*

Candidate genes validated for change in PAS selection and steady-state levels by qRT-PCR. Altered PAS selection was calculated as ΔΔCt of distal primer PCR product relative to total transcripts as detected by coding sequence primers and are reported in *Pabpn1*$^{+/A17}$ mice as a fraction of wild type. *D)* Steady-state levels of candidate genes as analyzed by qRT-PCR. All transcripts were normalized to *Gapdh*. *$p<0.05$. *E)* Venn diagram depicting changed steady-state mRNAs in RF from three-month-old *Pabpn1*$^{+/A17}$. Of 4515 genes detected, 95 were depleted and 30 were increased. *F)* Sample clustering showing that based on steady-state mRNA levels, *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice group by genotype. *G)* Histogram of the log P value of top GO terms and KEGG pathways for mRNAs decreased in *Pabpn1*$^{+/A17}$ mice revealing decreased levels of nuclear-encoded mitochondrial mRNAs.

Figure 2.5: Proteomic analysis of *Pabpn1*$^{+/A17}$ mice reveals altered mitochondrial proteins

A) Diagram depicting number of total proteins detected in rectus femoris from three-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice and those enriched or depleted in *Pabpn1*$^{+/A17}$ mice. *B)* Histogram of the log P value of top GO terms and KEGG pathways for proteins depleted from *Pabpn1*$^{+/A17}$ mice revealing depletion of mitochondrial proteins. *C)* Histogram of the log P value of top GO terms and KEGG pathways for proteins enriched in *Pabpn1*$^{+/A17}$ mice revealing increased levels of proteasome and cytoskeletal organization proteins. *D)* Immunoblot (left) from homogenized rectus femoris muscles from three-month-old *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice showing decreased levels of electron transport chain proteins (NDUFB8, SDHB, UQCR2, MT-

COX1, ATP5A) as quantified relative to total HSP-90 (right). *E)* No change in total mitochondrial content as detected by immunoblot for VDAC (top) and quantified relative to HSP-90 (bottom) in muscles from *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. *F)* Increased mitochondrial DNA (mtDNA) as detected by quantitative PCR of mtDNA markers *mtCox1* and *mtCox3* relative to nuclear genome markers *Hprt* and *Dio3* in total genomic DNA from *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/A17}$ mice. All quantitative data are mean ± SEM from n = 3 animals per genotype. **p<0.05.*

Figure 2.6: Age-related hypertrophy in *Pabpn1*$^{+/\Delta}$ knock-out mice

A) Schematic of *Pabpn1* knock-out allele. Before Cre recombinase-mediated recombination, the targeted allele contains a puromycin resistance gene flanked by *FRT* elements between exons 2 and 3 of the *Pabpn1* locus. The FRT-flanked puromycin resistance gene and exons 1 and 2 of Pabpn1 are surrounded by *loxP* sites. After flippase and Cre-mediated recombination, exons 1 and 2 of *Pabpn1* are excised resulting in a null allele with no start codon. *B)* Steady-state *Pabpn1* mRNA levels were decreased by ~50% in muscle and spleen from 3-month-old male *Pabpn1*$^{+/\Delta}$ mice as detected by qRT-PCR. Data are mean fold change ± SEM from n = 3 mice per genotype. **p<0.05. C)* Steady-state PABPN1 protein levels were decreased by half in 3-month-old male *Pabpn1*$^{+/\Delta}$ mice as detected by immunoblot using an α-PABPN1 antibody (50). HSP-90 was used as a loading control. *D)* Histologic sections from rectus femoris (RF) muscles isolated from six-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice shown at low and high magnification. Bar = 50 μm. *E)* Significantly smaller myofiber cross-sectional area (CSA) was detected in RF muscles from six-month-old *Pabpn1*$^{+/\Delta}$ mice. Data are the

frequency distribution of myofiber CSA. *F)* No significant difference detected in the percentage of centrally nucleated fibers (CNF) from RF sections from six-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice. Data are mean ± SEM. *G)* Histologic sections from RF muscle isolated from 18-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice shown at low and high magnification. Bar = 50 μm. *H)* Significantly larger CSA was detected in RF muscles from 18-month-old male *Pabpn1*$^{+/\Delta}$ mice. Data are the frequency distribution of myofiber CSA. *I)* No significant difference was detected in the percentage of centrally nucleated fibers (CNF) from RF sections from 18-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice. Data are mean ± SEM. In all cases, n = 5 mice per genotype.

Figure 2.7: Decreased poly(A) tail length but no change in PAS usage in *Pabpn1*$^{+/\Delta}$ mice

A) Poly(A) tail length was analyzed by polyacrylamide gel electrophoresis of 3' end labeled poly(A) tracts from total RNA isolated from rectus femoris muscles of three-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice. n = 3 per genotype. *B*) Quantification of regions representing short (1), medium (2), and long (3) poly(A) tails noted in (*A*) normalized to the total short poly(A) tails marked in (*A*) (Normalize) revealed fewer long poly(A) tails in *Pabpn1*$^{+/A17}$ mice. *C*) Quantification of gel from *(A)* by plotting density along each lane revealed shorter poly(A) tails in *Pabpn1*$^{+/\Delta}$ mice. ****$p < 0.0001$ *D)* Scatter plot depicting few changes were detected in poly(A) sequence (PAS) selection as detected by 3'READS. Data are $\log_2$ ratio of distal to proximal PAS usage with increased distal PAS represented in red and increased proximal PAS represented in blue. *E)* Sample

clustering based on PAS usage as detected by 3'READS showing that *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice do not group by genotype. n = 2 mice per genotype.

Figure 2.8: Mitochondrial defects detected in *Pabpn1*$^{+/\Delta}$ mice

A) Diagram depicting number of total proteins detected in rectus femoris muscles from three-month-old male *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice and those enriched or depleted in *Pabpn1*$^{+/\Delta}$ mice. *B)* Histogram of the log P value of top GO terms and KEGG pathways for proteins depleted from *Pabpn1*$^{+/\Delta}$ mice revealing decreased levels of mitochondrial proteins. Groups that are also depleted in *Pabpn1*$^{+/A17}$ mice are marked with a green asterisk. *C)* Histogram of the log p value of top GO terms and KEGG pathways for proteins enriched in *Pabpn1*$^{+/\Delta}$ mice revealing increased levels of mitochondrial and proteasome proteins. Groups that are depleted rather than enriched in *Pabpn1*$^{+/A17}$ mice are marked with a red asterisk. *D)* Increased levels of electron transport chain proteins (SDHB, UQCR2, ATP5A) as detected by immunoblot (left) and quantified relative to total HSP-90 (*E*) in *Pabpn1*$^{+/\Delta}$ mice. *F)* No change in total mitochondrial content as detected by immunoblot for VDAC (top) and quantified relative to HSP-90 (bottom) in *Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice. *G)* Decreased mitochondrial DNA (mtDNA) as detected by quantitative PCR of mtDNA markers *mtCox1* and *mtCox3* relative to nuclear genome markers *Hprt* and *Dio3* in total genomic DNA from*Pabpn1*$^{+/+}$ and *Pabpn1*$^{+/\Delta}$ mice. All quantitative data are mean ± SEM from n = 3 animals per genotype. *$p<0.05$.

Supplemental Figures are available through the journal Human Molecular Genetics.

Figure 2.1

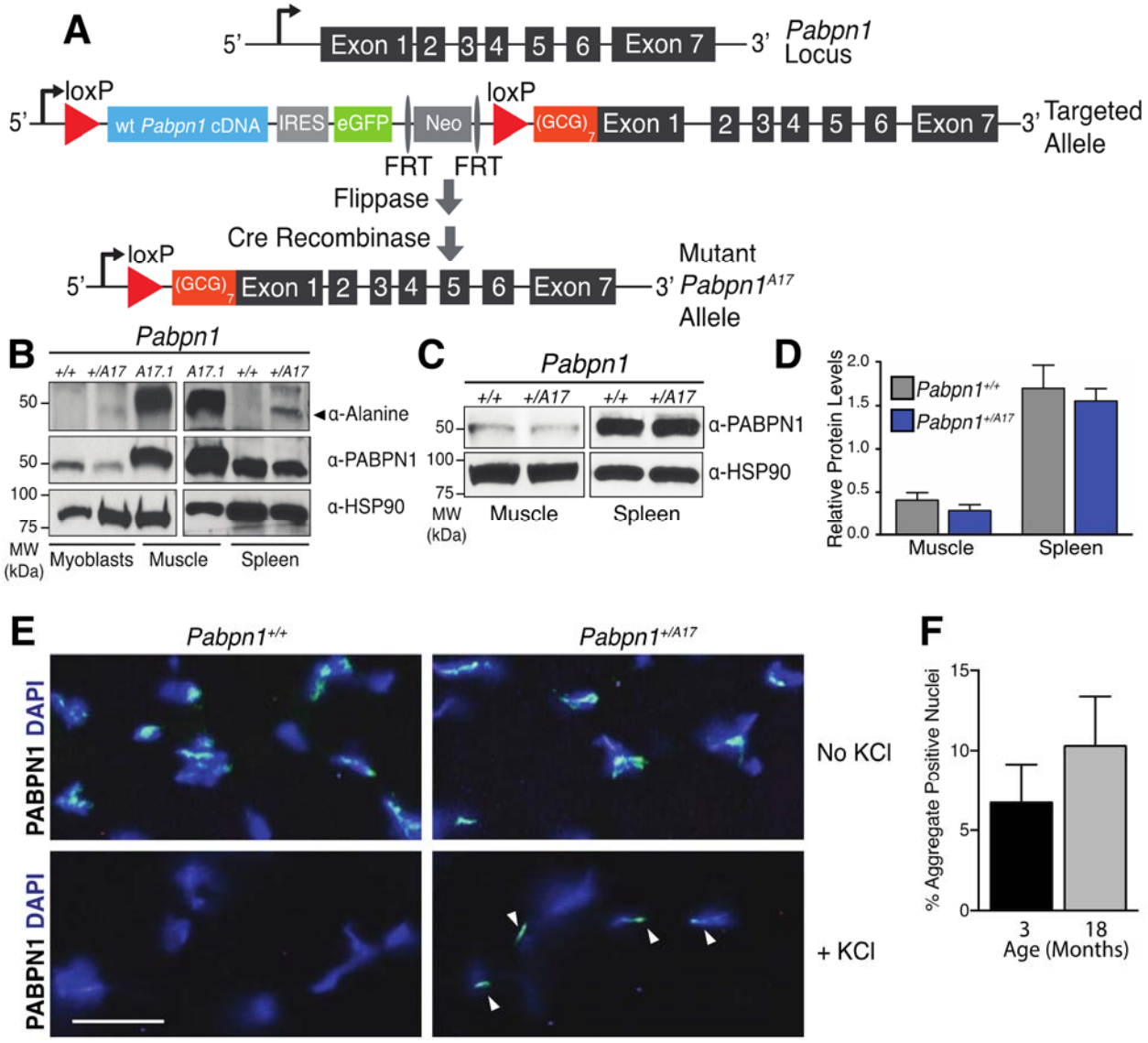

A
5'
Exon 1 2 3 4 5 6 Exon 7
3' Pabpn1 Locus
loxP
wt Pabpn1 cDNA
IRES
eGFP
Neo
FRT FRT
loxP
(GCG)7 Exon 1 2 3 4 5 6 Exon 7
3' Targeted Allele
Flippase
Cre Recombinase
loxP
(GCG)7 Exon 1 2 3 4 5 6 Exon 7
3' Pabpn1A17 Allele
Mutant
B
Pabpn1
+/+ +/A17 A17.1 A17.1 +/+ +/A17
50
50
100
75
MW (kDa)
α-Alanine
α-PABPN1
α-HSP90
Myoblasts
Muscle
Spleen
C
Pabpn1
+/+ +/A17 +/+ +/A17
50
100
75
MW (kDa)
α-PABPN1
α-HSP90
Muscle
Spleen
D
Relative Protein Levels
2.0
1.5
1.0
0.5
0.0
Pabpn1+/+
Pabpn1+/A17
Muscle
Spleen
E
Pabpn1+/+
Pabpn1+/A17
PABPN1 DAPI
PABPN1 DAPI
No KCl
+ KCl
F
% Aggregate Positive Nuclei
15
10
5
0
3
18
Age (Months)

Figure 2.2

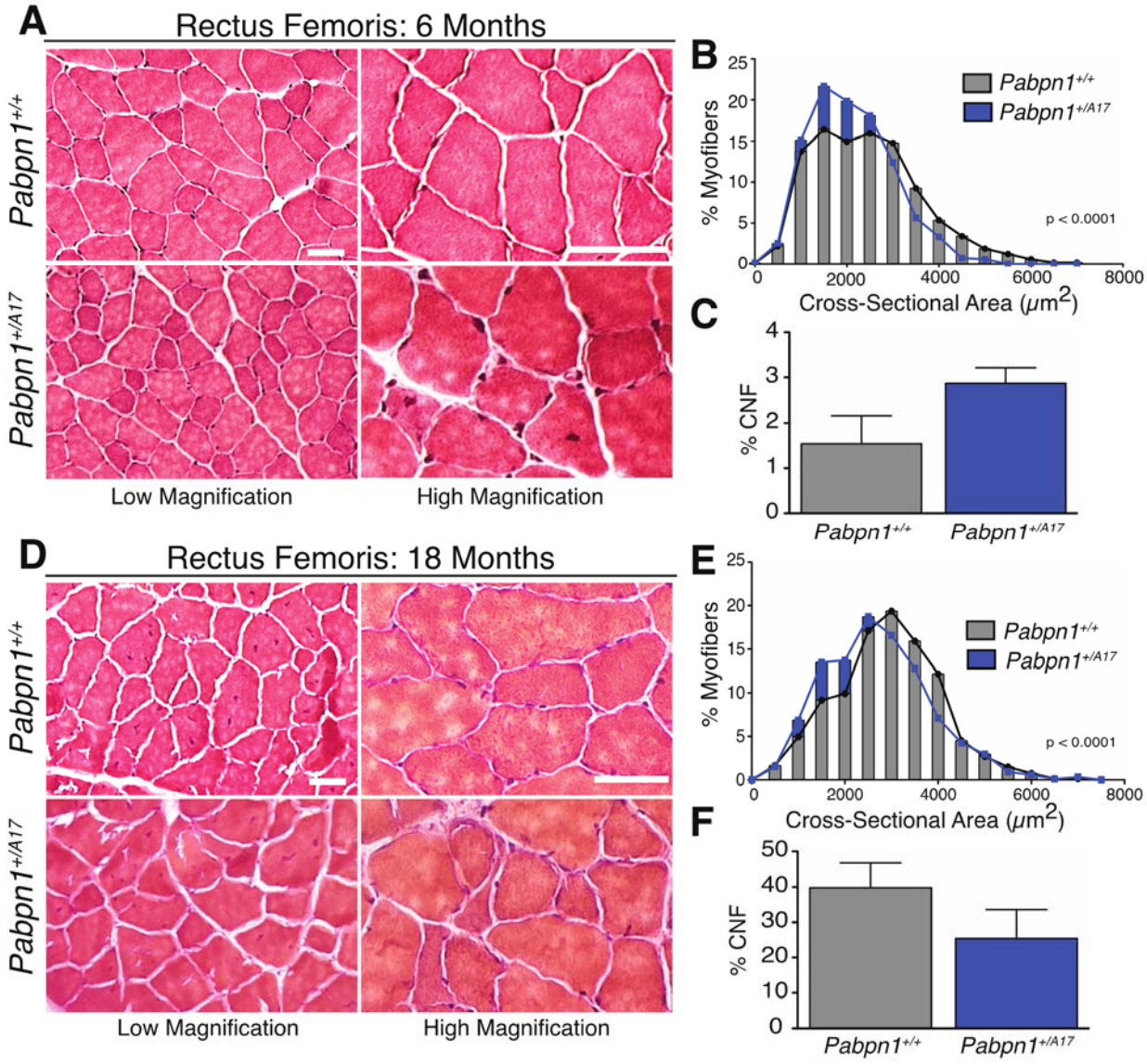

A
Rectus Femoris: 6 Months
Pabpn1+/+
Pabpn1+/A17
Low Magnification
High Magnification
B
% Myofibers
Pabpn1+/+
Pabpn1+/A17
p < 0.0001
Cross-Sectional Area (μm2)
C
% CNF
Pabpn1+/+
Pabpn1+/A17
D
Rectus Femoris: 18 Months
Pabpn1+/+
Pabpn1+/A17
Low Magnification
High Magnification
E
% Myofibers
Pabpn1+/+
Pabpn1+/A17
p < 0.0001
Cross-Sectional Area (μm2)
F
% CNF
Pabpn1+/+
Pabpn1+/A17

Figure 2.3

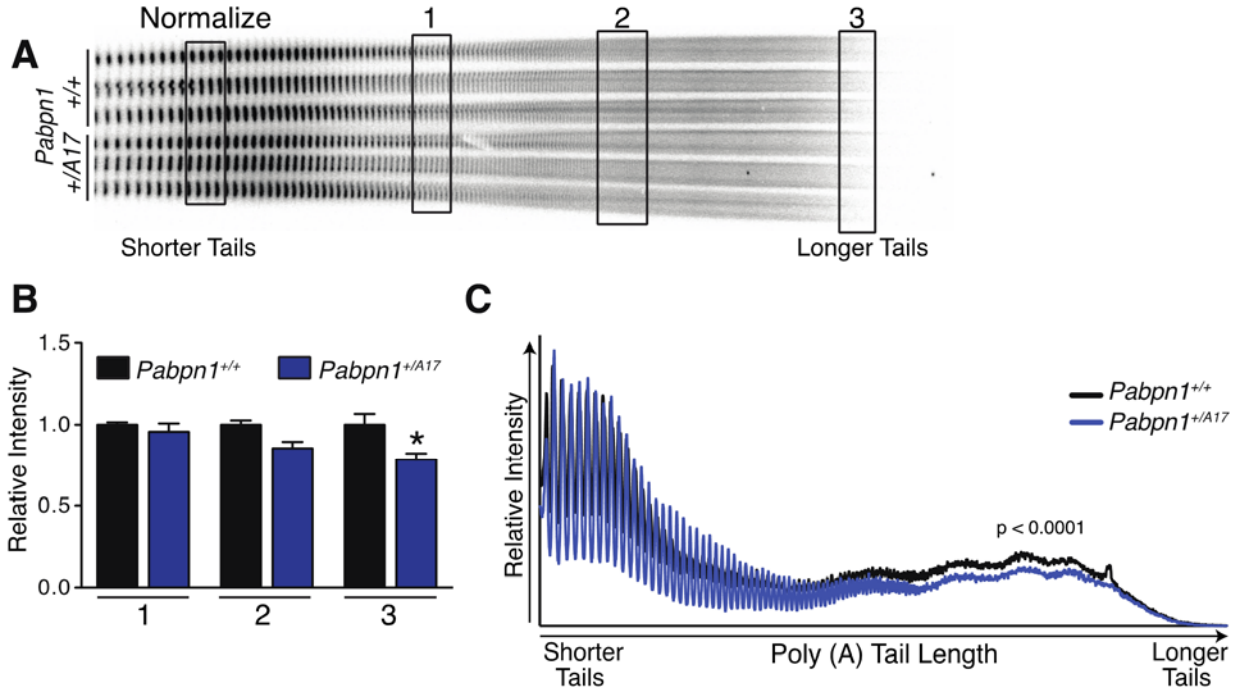

A
Normalize
1
2
3
Pabpn1
+/+
+/A17
Shorter Tails
Longer Tails
B
Relative Intensity
1.5
1.0
0.5
0.0
Pabpn1+/+
Pabpn1+/A17
*
1
2
3
C
Relative Intensity
Pabpn1+/+
Pabpn1+/A17
p < 0.0001
Shorter Tails
Poly (A) Tail Length
Longer Tails

Figure 2.4

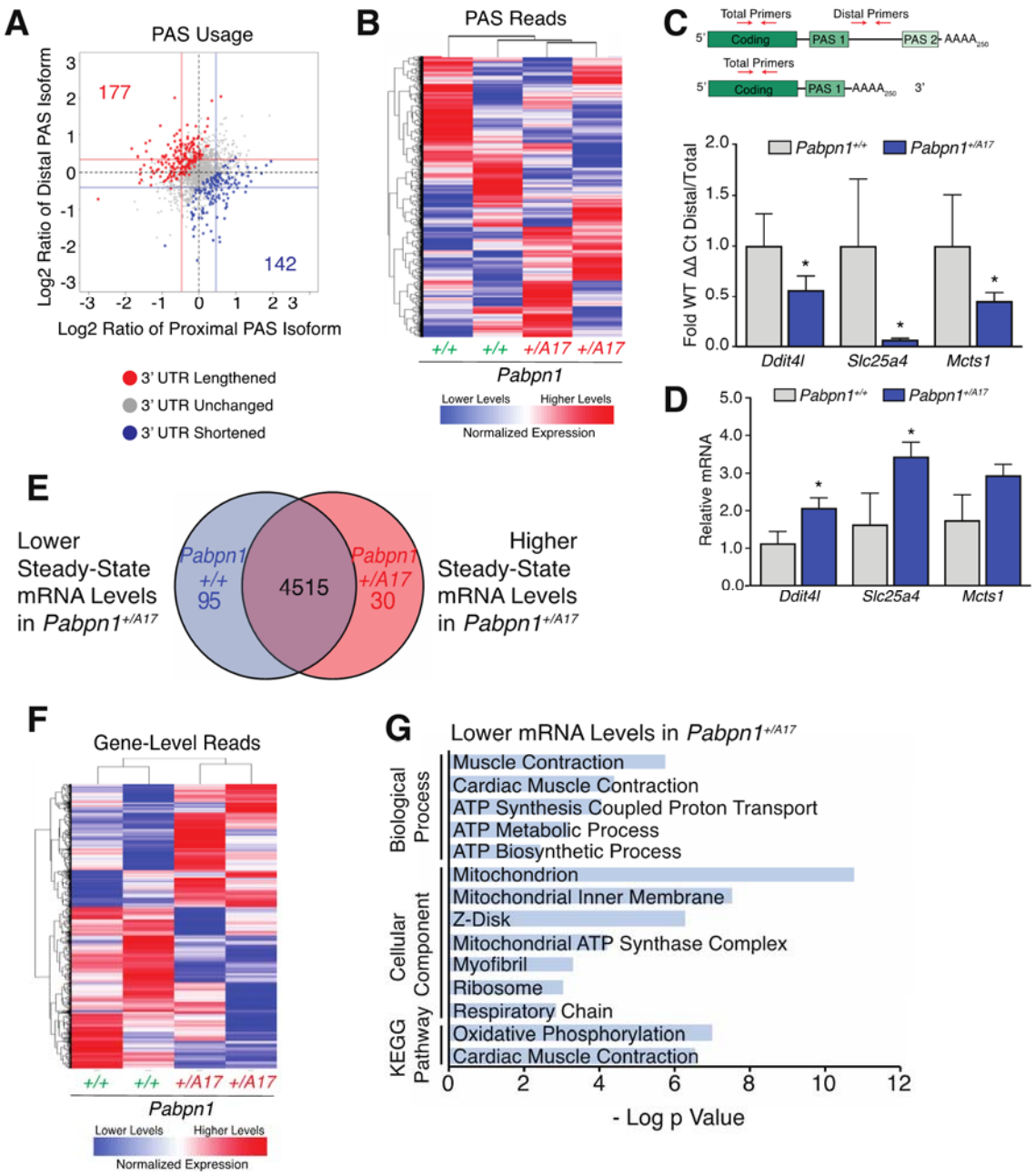

A
PAS Usage
Log2 Ratio of Distal PAS Isoform
Log2 Ratio of Proximal PAS Isoform
177
142
3' UTR Lengthened
3' UTR Unchanged
3' UTR Shortened
B
PAS Reads
+/+ +/+ +/A17 +/A17
Pabpn1
Lower Levels
Higher Levels
Normalized Expression
C
Total Primers
Distal Primers
Coding
PAS 1
PAS 2
AAAA
Fold WT ΔΔ Ct Distal/Total
Pabpn1+/+
Pabpn1+/A17
Ddit4l
Slc25a4
Mcts1
D
Relative mRNA
E
Lower Steady-State mRNA Levels in Pabpn1+/A17
Pabpn1 +/+ 95
4515
Pabpn1 +/A17 30
Higher Steady-State mRNA Levels in Pabpn1+/A17
F
Gene-Level Reads
G
Lower mRNA Levels in Pabpn1+/A17
Biological Process
Muscle Contraction
Cardiac Muscle Contraction
ATP Synthesis Coupled Proton Transport
ATP Metabolic Process
ATP Biosynthetic Process
Cellular Component
Mitochondrion
Mitochondrial Inner Membrane
Z-Disk
Mitochondrial ATP Synthase Complex
Myofibril
Ribosome
KEGG Pathway
Respiratory Chain
Oxidative Phosphorylation
Cardiac Muscle Contraction
- Log p Value

Figure 2.5

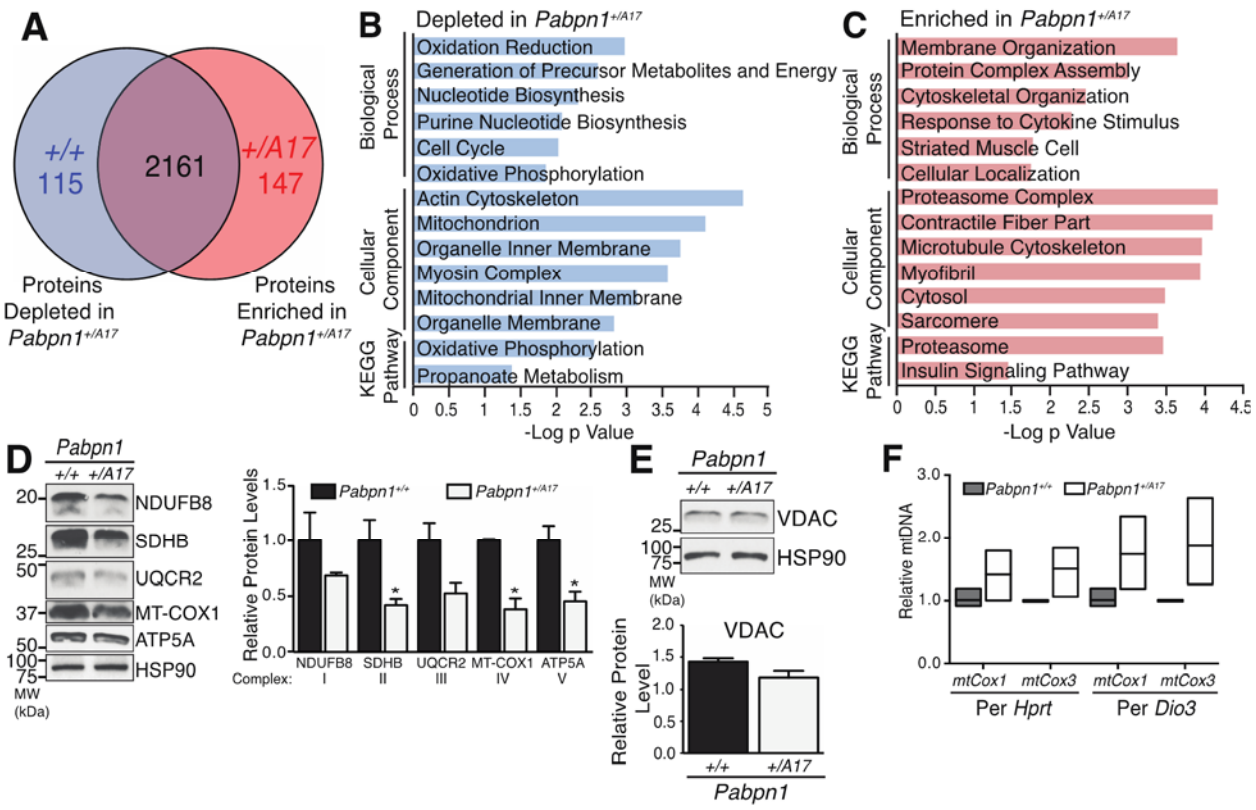

A
+/+
115
2161
+/A17
147
Proteins Depleted in Pabpn1+/A17
Proteins Enriched in Pabpn1+/A17
B
Depleted in Pabpn1+/A17
Biological Process
Oxidation Reduction
Generation of Precursor Metabolites and Energy
Nucleotide Biosynthesis
Purine Nucleotide Biosynthesis
Cell Cycle
Oxidative Phosphorylation
Cellular Component
Actin Cytoskeleton
Mitochondrion
Organelle Inner Membrane
Myosin Complex
Mitochondrial Inner Membrane
Organelle Membrane
KEGG Pathway
Oxidative Phosphorylation
Propanoate Metabolism
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5
-Log p Value
C
Enriched in Pabpn1+/A17
Biological Process
Membrane Organization
Protein Complex Assembly
Cytoskeletal Organization
Response to Cytokine Stimulus
Striated Muscle Cell
Cellular Localization
Cellular Component
Proteasome Complex
Contractile Fiber Part
Microtubule Cytoskeleton
Myofibril
Cytosol
Sarcomere
KEGG Pathway
Proteasome
Insulin Signaling Pathway
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
-Log p Value
D
Pabpn1
+/+ +/A17
20
NDUFB8
SDHB
25
50
UQCR2
37
MT-COX1
50
ATP5A
100
75
HSP90
MW (kDa)
Pabpn1+/+
Pabpn1+/A17
Relative Protein Levels
1.5
1.0
0.5
0.0
NDUFB8 SDHB UQCR2 MT-COX1 ATP5A
Complex: I II III IV V
E
Pabpn1
+/+ +/A17
25
VDAC
100
75
HSP90
MW (kDa)
VDAC
Relative Protein Level
2.0
1.5
1.0
0.5
0.0
+/+ +/A17
Pabpn1
F
Pabpn1+/+
Pabpn1+/A17
Relative mtDNA
3.0
2.0
1.0
0.0
mtCox1 mtCox3 mtCox1 mtCox3
Per Hprt
Per Dio3

Figure 2.6

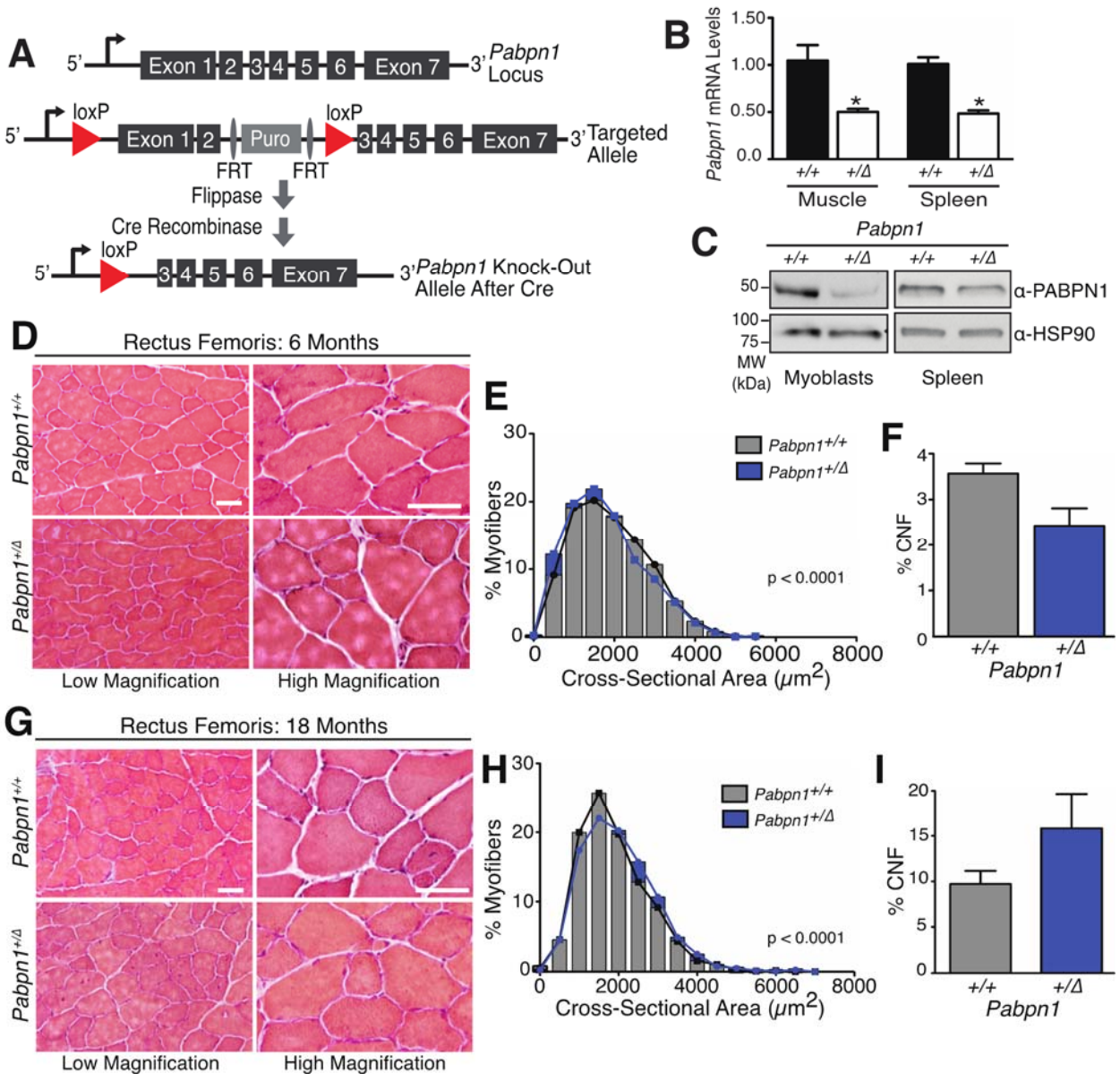

A
5'
Exon 1 2 3 4 5 6 Exon 7
3' Pabpn1 Locus
loxP
Puro
FRT
FRT
3' Targeted Allele
Flippase
Cre Recombinase
3' Pabpn1 Knock-Out Allele After Cre
B
Pabpn1 mRNA Levels
1.50
1.00
0.50
0.0
+/+ +/Δ +/+ +/Δ
Muscle
Spleen
C
Pabpn1
+/+ +/Δ +/+ +/Δ
50
100
75
MW (kDa)
α-PABPN1
α-HSP90
Myoblasts
Spleen
D
Rectus Femoris: 6 Months
Pabpn1+/+
Pabpn1+/Δ
Low Magnification
High Magnification
E
% Myofibers
Pabpn1+/+
Pabpn1+/Δ
p < 0.0001
Cross-Sectional Area (μm2)
F
% CNF
+/+ +/Δ
Pabpn1
G
Rectus Femoris: 18 Months
Pabpn1+/+
Pabpn1+/Δ
Low Magnification
High Magnification
H
% Myofibers
Pabpn1+/+
Pabpn1+/Δ
p < 0.0001
Cross-Sectional Area (μm2)
I
% CNF
+/+ +/Δ
Pabpn1

Figure 2.7

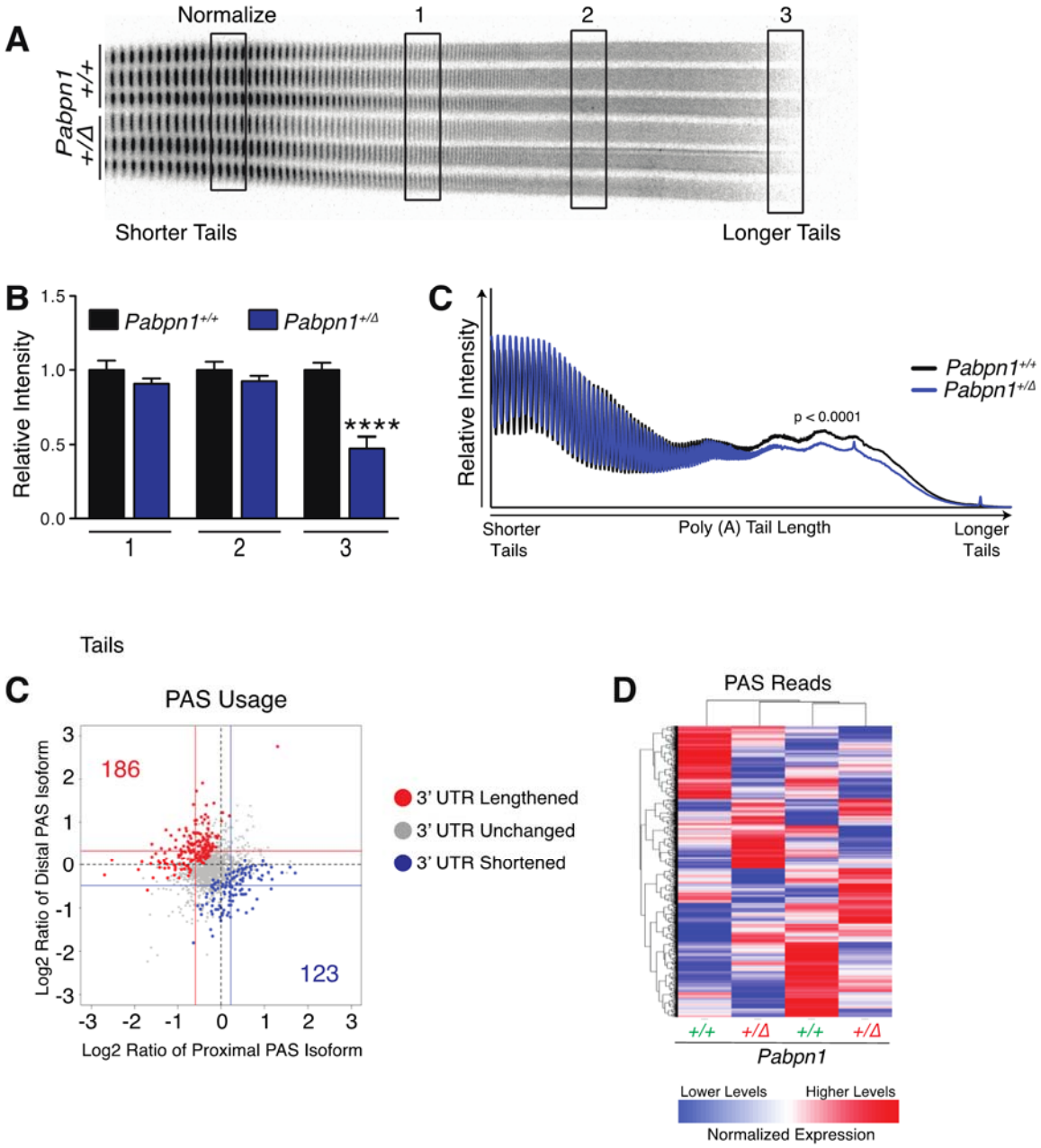
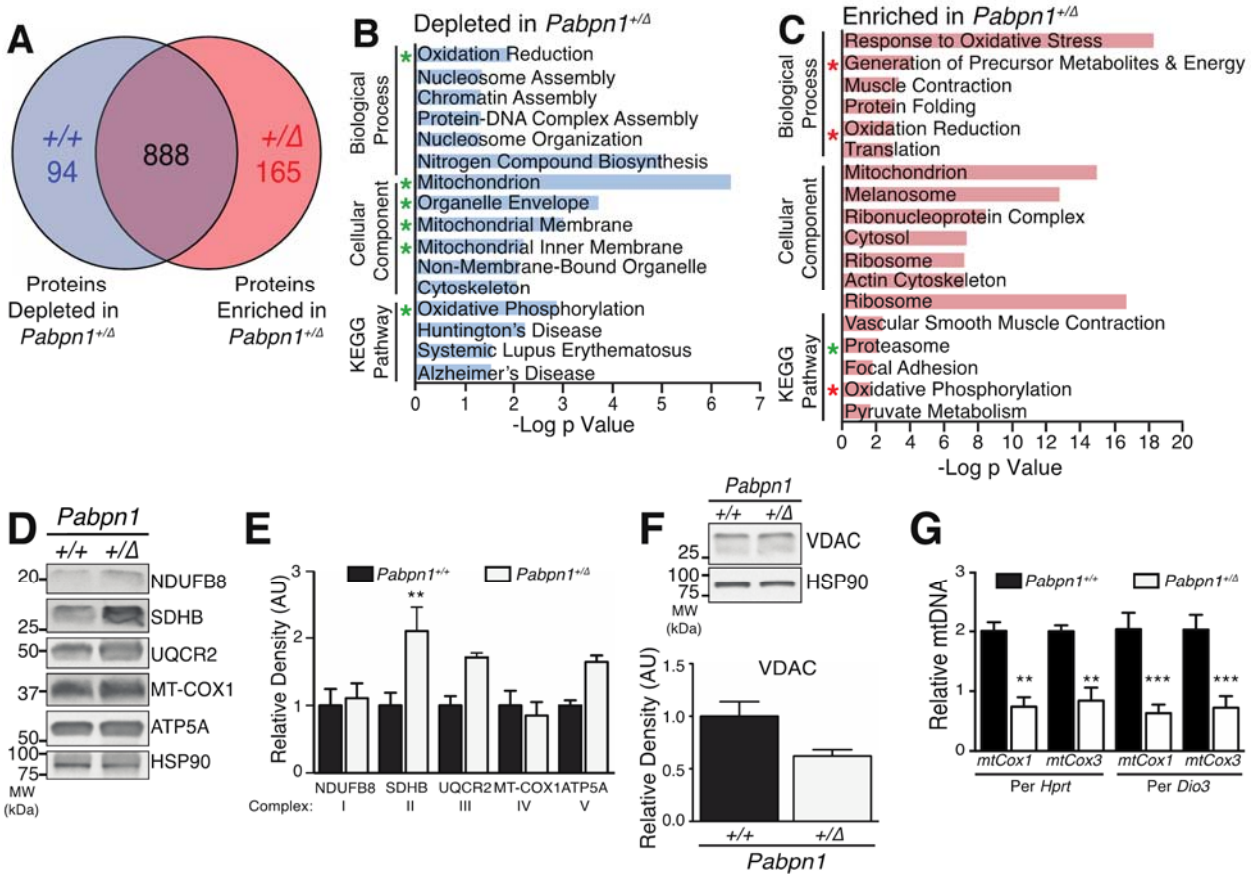

A
Normalize
1
2
3
Pabpn1 +/+
Pabpn1 +/Δ
Shorter Tails
Longer Tails
B
Pabpn1+/+
Pabpn1+/Δ
Relative Intensity
1.5
1.0
0.5
0.0
****
1
2
3
C
Relative Intensity
Pabpn1+/+
Pabpn1+/Δ
p < 0.0001
Shorter Tails
Poly (A) Tail Length
Longer Tails
Tails
C
PAS Usage
186
123
Log2 Ratio of Distal PAS Isoform
Log2 Ratio of Proximal PAS Isoform
3' UTR Lengthened
3' UTR Unchanged
3' UTR Shortened
D
PAS Reads
+/+
+/Δ
+/+
+/Δ
Pabpn1
Lower Levels
Higher Levels
Normalized Expression

Figure 2.8


A
+/+
94
888
+/Δ
165
Proteins Depleted in Pabpn1+/Δ
Proteins Enriched in Pabpn1+/Δ
B
Depleted in Pabpn1+/Δ
Biological Process
Oxidation Reduction
Nucleosome Assembly
Chromatin Assembly
Protein-DNA Complex Assembly
Nucleosome Organization
Nitrogen Compound Biosynthesis
Cellular Component
Mitochondrion
Organelle Envelope
Mitochondrial Membrane
Mitochondrial Inner Membrane
Non-Membrane-Bound Organelle
Cytoskeleton
KEGG Pathway
Oxidative Phosphorylation
Huntington's Disease
Systemic Lupus Erythematosus
Alzheimer's Disease
0 1 2 3 4 5 6 7
-Log p Value
C
Enriched in Pabpn1+/Δ
Biological Process
Response to Oxidative Stress
Generation of Precursor Metabolites & Energy
Muscle Contraction
Protein Folding
Oxidation Reduction
Translation
Cellular Component
Mitochondrion
Melanosome
Ribonucleoprotein Complex
Cytosol
Ribosome
Actin Cytoskeleton
Ribosome
KEGG Pathway
Vascular Smooth Muscle Contraction
Proteasome
Focal Adhesion
Oxidative Phosphorylation
Pyruvate Metabolism
0 2 4 6 8 10 12 14 16 18 20
-Log p Value
D
Pabpn1
+/+ +/Δ
20
NDUFB8
25
SDHB
50
UQCR2
37
MT-COX1
50
ATP5A
100
75
HSP90
MW (kDa)
E
Pabpn1+/+
Pabpn1+/Δ
Relative Density (AU)
**
NDUFB8 SDHB UQCR2 MT-COX1 ATP5A
Complex: I II III IV V
F
Pabpn1
+/+ +/Δ
25
VDAC
100
75
HSP90
MW (kDa)
VDAC
Relative Density (AU)
+/+ +/Δ
Pabpn1
G
Pabpn1+/+
Pabpn1+/Δ
Relative mtDNA
**
**
***
***
mtCox1 mtCox3 mtCox1 mtCox3
Per Hprt
Per Dio3

Chapter 3: Post-transcriptional regulation of *Pabpn1* by the RNA binding protein HuR

A portion of this chapter is under revision at the journal Nucleic Acids Research:

Phillips B.L., Banerjee A., Sanchez B.J., Di Marco S., Gallouzi I.E., Pavlath G.K., Corbett A.H. *Post-transcriptional regulation of Pabpn1 by the RNA binding protein HuR.* (2018) *In revision.*

Brittany Phillips performed experiments and data analyses for all figures (Figures 3.1-3.8 and Supplementary Figures 3.1-3.8) with the exception of Figure 3.4A and 3.4B, which were performed by Ayan Banerjee.

3.1 Summary

RNA processing is critical for proper spatial and temporal control of gene expression. The ubiquitous nuclear polyadenosine RNA binding protein, PABPN1, post-transcriptionally regulates multiple steps of gene expression. Mutations in the *PABPN1* gene that expand an N-terminal alanine tract in the PABPN1 protein from 10 alanines to 11-18 alanines cause the muscle-specific disease oculopharyngeal muscular dystrophy (OPMD), which affects muscles of the eyelid, pharynx, and proximal limbs. Previous work revealed that the *Pabpn1* transcript is unstable, contributing to low steady-state *Pabpn1* mRNA and protein levels *in vivo*, specifically in skeletal muscle, with even lower levels in muscles affected in OPMD. Thus, low levels of PABPN1 protein could predispose specific tissues to pathology in OPMD. However, no studies have defined the mechanisms that regulate *Pabpn1* expression. Here, we define multiple *cis*-regulatory elements and a *trans*-acting factor, HuR, which regulate *Pabpn1* expression specifically in mature muscle *in vitro* and *in vivo*. We exploit multiple models including C2C12 myotubes, primary muscle cells, and mice to determine that HuR regulates *Pabpn1* expression. Overall, we have uncovered a mechanism specific to mature muscle that regulates *Pabpn1* expression *in vitro* and *in vivo*, which could provide insight to future studies investigating therapeutic strategies for OPMD treatment.

3.2 Introduction

RNA processing ensures proper spatial and temporal control of gene expression. RNA binding proteins regulate every aspect of RNA processing from 5’ capping and 3’

end formation in the nucleus to eventual RNA turnover in the cytoplasm. Consistent with the critical roles for RNA binding proteins in gene expression, the vast majority of these proteins are ubiquitously expressed (129). Despite this ubiquitous requirement for RNA binding proteins, mutations in genes encoding RNA binding proteins often cause tissue-specific disease (36). The molecular deficiencies underlying tissue-specificity are typically difficult to define and likely depend on the functions of the RNA binding protein as well as the context and specific requirements within the affected tissue. Interestingly, mutations in genes encoding RNA binding proteins often affect nervous and muscle tissue (36), suggesting that these complex tissues have strict requirements for proper RNA processing and control of gene expression. One example of an RNA binding protein linked to a tissue-specific disease is the nuclear polyadenosine (poly[A]) binding protein PABPN1. Mutations in the ubiquitously-expressed *PABPN1* gene cause the muscle-specific disease oculopharyngeal muscular dystrophy (OPMD) (38).

Although PABPN1 regulates multiple aspects of RNA processing (33), the best characterized role for PABPN1 is controlling poly(A) tail length by stimulating poly(A) polymerase (46,47). More recent work has defined how PABPN1 and polyadenylation serve as quality control mechanisms. Depletion of *Pabpn1* causes global poly(A) tail shortening and nuclear poly(A) RNA retention in primary muscle cells, demonstrating that PABPN1-mediated RNA processing events are necessary for efficient RNA nuclear export (50). Furthermore, PABPN1 and poly(A) polymerase hyperadenylate improperly spliced or intron-retaining RNAs, targeting these RNAs for exosome-mediated degradation (40). PABPN1 also modulates alternative polyadenylation (2,41,42,44,130), demonstrating that PABPN1 regulates 3’ untranslated region (3’UTR) length which could

influence downstream post-transcriptional regulatory mechanisms. Thus, PABPN1 plays critical roles in numerous RNA processing events important for proper gene expression.

In addition to a conventional RNA recognition motif (RRM) (45) that mediates RNA binding, the PABPN1 protein contains a 10-alanine tract immediately following the initial methionine (38) in a region of the protein with no characterized function. Patients with OPMD have GCN triplet expansion mutations that extend this 10-alanine tract to 11-18 alanines (38). How this subtle expansion in a region of the PABPN1 protein with no known function causes a dominant disease affecting a specific subset of muscles in the eyelid, pharynx, and proximal limbs is not understood.

Studies aimed at understanding how this modest alanine expansion in PABPN1 confers pathology in a subset of skeletal muscles have typically focused on the propensity of alanine-expanded PABPN1 to form insoluble nuclear aggregates (2,73). Aggregate-positive myonuclei from a transgenic mouse model overexpressing expanded PABPN1 are positive for cell death markers (56), suggesting that this alanine tract expansion may induce toxicity. However, both wildtype and expanded PABPN1 can aggregate *in vitro,* and this aggregation does not require the N-terminus containing the alanine tract (57). Furthermore, nuclear aggregates are only present in ~5% of myonuclei (73) and also occur in the brain (78), a tissue that is not typically affected in patients with OPMD (33). These observations raise questions about whether aggregate-driven toxicity is the sole mechanism of pathology in OPMD.

Aggregate-mediated disease pathology would logically occur in tissues with high levels of the mutant protein. If toxic PABPN1 aggregates were the sole source of pathology in OPMD, tissues that express high levels of the toxic protein would likely be

the most susceptible to pathology. However, PABPN1 levels are extremely low in skeletal muscles compared to other tissues not affected in OPMD (58). In fact, PABPN1 protein levels are even lower in muscles that are typically affected in OPMD patients such as pharynx than other skeletal muscles (58). These low PABPN1 levels may predispose particular muscles to OPMD pathology.

As PABPN1 levels are low in muscle, a reduction in the functional pool of PABPN1 protein may more severely affect muscle than tissues with higher PABPN1 levels. Thus, in OPMD, functional PABPN1 levels may fall below a threshold for pathology only in muscle while non-muscle tissues have sufficient levels of functional PABPN1 to fulfill tissue-specific requirements (58). Immunofluorescence studies have identified PABPN1 as well as other proteins sequestered in insoluble aggregates in OPMD patient muscle tissue (60,61), and overexpression studies have identified specific RNAs present in these aggregates (62). Thus, aggregation could cause or contribute to a decrease in the functional pool of PABPN1 in muscle where levels of PABPN1 are already low compared to other tissues. Together, these observations suggest a more complex disease pathology than purely aggregate-mediated toxicity. Therefore, increasing PABPN1 protein levels in muscles typically affected by PABPN1 expansion could be protective and considered as a component of future OPMD therapies. Unfortunately, little is known about any mechanisms that regulate expression of *PABPN1*.

The bulk of muscle-specific regulation of PABPN1 levels occurs at the RNA level as steady-state *Pabpn1* mRNA levels are low in muscle *in vivo* (58). Furthermore, the *Pabpn1* transcript is highly unstable in muscle relative to non-muscle tissue *in vivo*,

demonstrating that *Pabpn1* transcript instability contributes to low PABPN1 protein levels that are observed in muscle (58). Defining the mechanisms that regulate the *Pabpn1* transcript in muscle may identify novel targets that modulate PABPN1 protein levels.

To interrogate the specific mechanisms regulating the *Pabpn1* transcript, we combined the strengths of an *in vitro* model of mature skeletal muscle, C2C12 myotubes, with *in vivo* analyses. Here we validate C2C12 myotubes as an *in vitro* model of skeletal muscle in which to define the mechanisms that regulate *Pabpn1* in muscle. We define *cis*-regulatory elements in the *Pabpn1* 3'UTR including an AU-rich element (ARE) bound by the RNA binding protein HuR. We identify HuR as a direct regulator of *Pabpn1* transcript and protein levels. Furthermore, HuR-mediated *Pabpn1* regulation is conserved *in vivo* in muscles and primary muscle cells. These data demonstrate that HuR negatively regulates *Pabpn1* at the RNA and protein levels, providing insight into the mechanisms regulating *Pabpn1* expression in a mature skeletal muscle-specific manner. This regulatory mechanism could be exploited as a novel therapeutic approach to increase PABPN1 protein levels in OPMD.

3.3 Results

3.3.1 Steady-state *Pabpn1* mRNA and protein levels are low in an *in vitro* model of skeletal muscle

Our previous studies revealed that *Pabpn1* mRNA and protein is present at very low levels in both mouse and human skeletal muscle tissue (58). However, investigating the mechanisms underlying this regulation of *Pabpn1* expression in tissue is challenging. For example, identifying and defining *cis*-elements and *trans*-acting factors that influence *Pabpn1* expression *in vivo* would require generation of multiple mouse models. Therefore, identifying an appropriate and tractable *in vitro* cell culture model is necessary to dissect these skeletal muscle-specific mechanisms regulating *Pabpn1* expression.

Previous studies show that proliferative muscle progenitor cells called myoblasts, have higher PABPN1 levels than mature skeletal muscle (58), similar to levels observed in non-muscle tissue (58). Serum deprivation induces myoblasts to form terminally differentiated, multinucleated myotubes, a well-characterized *in vitro* model of mature skeletal muscle. Therefore, to explore if myoblasts and myotubes are appropriate for studying *Pabpn1* regulation *in vitro*, we employed the immortalized mouse myoblast cell line C2C12 (131).

We differentiated C2C12 myoblasts (Figure 3.1A) into multinucleated myotubes and used morphological (Figure 3.1B) and biochemical analyses (Figure 3.1C) to confirm efficient differentiation. The late differentiation marker embryonic Myosin Heavy Chain (eMyHC) (132) is readily detected in myotubes but is undetectable in undifferentiated myoblasts (Figure 3.1C, Supplementary Figure 1).

We then compared PABPN1 protein levels in C2C12 myoblasts and C2C12 myotubes to determine whether this *in vitro* system recapitulates patterns of PABPN1 expression observed *in vivo* (58). As shown in Figure 3.1C,D, PABPN1 protein levels decrease ~40% in C2C12 myotubes relative to C2C12 myoblasts. To investigate whether this myotube-specific reduction in PABPN1 protein levels occurs at the RNA level, we used qRT-PCR to assess steady-state *Pabpn1* transcript levels. Steady-state *Pabpn1* mRNA levels are reduced ~50% in C2C12 myotubes compared to C2C12 myoblasts (Figure 3.1E), suggesting that the bulk of differential PABPN1 regulation occurs at the level of RNA. These data are consistent with *in vivo* data showing that *Pabpn1* mRNA and protein levels are low in skeletal muscle relative to primary murine myoblasts and non-muscle tissue (58), demonstrating that this *in vitro* system is appropriate for studying the mechanisms that regulate *Pabpn1* expression in mature muscle.

3.3.2 The *Pabpn1* transcript is unstable in an *in vitro* model of skeletal muscle

In vivo studies identified decreased mRNA stability as one mechanism regulating *Pabpn1* mRNA levels in skeletal muscle (58). We investigated whether *Pabpn1* mRNA stability differs between C2C12 myoblasts and myotubes as these myotubes are an *in vitro* model of mature skeletal muscle. To determine whether *Pabpn1* stability changes in a muscle cell- or myotube-specific manner, we also analyzed *Pabpn1* stability in the mouse fibroblast cell line NIH/3T3 (133). For these experiments, we treated cells with the transcriptional inhibitor Actinomycin D (134) and measured *Pabpn1* mRNA levels over time by qRT-PCR to determine the half-life of the *Pabpn1* transcript. As shown in Figure 3.2, *Pabpn1* mRNA is stable in both NIH/3T3 (Figure 3.2A) and C2C12

myoblasts (Figure 3.2B) over the time course examined. In contrast, the *Pabpn1* transcript is unstable in C2C12 myotubes (Figure 3.2C) relative to both NIH/3T3 and C2C12 myoblasts (Supplementary Figure 2A), with a projected half-life of ~13 h. As a control, *Myc,* a transcript known to be unstable in multiple cell lines and tissue types (135,136), is unstable with a $t_{1/2}$ of ~0.8 h in all cell lines analyzed (Figure 3.2, Supplementary Figure 2B). These data demonstrate that the *Pabpn1* transcript is less stable in C2C12 myotubes as compared to NIH/3T3 and C2C12 myoblast cells, illustrating that mechanisms regulating *Pabpn1* in muscle tissue can be recapitulated in C2C12 myotubes and are likely specific to mature muscle.

3.3.3 *Pabpn1* alternative polyadenylation does not correlate with stability changes

To further investigate myotube-specific *Pabpn1* regulation, we focused on the *Pabpn1* 3' untranslated region (3'UTR) as 3'UTRs often contain *cis*-regulatory elements that interact with *trans*-acting factors, contributing to spatial and temporal control of gene expression (20). The *Pabpn1* 3'UTR is highly conserved across several mammalian species with bovine and murine *Pabpn1* 3'UTRs sharing 96% and 93% sequence identity within the ~900 nucleotide human *Pabpn1* 3'UTR, respectively (Supplementary Figure 3A), suggesting that regulatory elements in the *Pabpn1* 3'UTR are likely also conserved.

The *Pabpn1* 3'UTR contains multiple predicted *cis*-regulatory elements, including two conserved cleavage/polyadenylation sites (PAS) (58) (Figure 3.3A). PAS utilization specifies where cleavage and polyadenylation occur which determines 3' end formation (137-139). If a transcript contains multiple PAS sites, the PAS utilized defines

3'UTR length and potentially inclusion or exclusion of regulatory elements with possible downstream effects on post-transcriptional regulation.

The predominant *Pabpn1* mRNA variant utilizes the distal PAS in all murine tissues examined *in vivo* with the exception of testis (58) To investigate whether PAS usage changes in C2C12 myotubes relative to C2C12 myoblasts, we conducted qRT-PCR using primers that amplify either the coding DNA sequence (CDS) or the distal region of the *Pabpn1* transcript downstream of PASI (Figure 3.3B). We found no significant change in the ratio of *Pabpn1* alternative PAS utilization comparing C2C12 myoblasts and C2C12 myotubes (Figure 3.3C), indicating that changes in inclusion or exclusion of regulatory elements does not likely drive differential regulation of *Pabpn1* stability between these cell types. To further investigate *Pabpn1* mRNA variants in C2C12 myoblasts and myotubes, we conducted northern blotting using a radiolabeled *Pabpn1* probe designed to detect *Pabpn1* transcripts generated from use of either PAS. Northern blotting confirmed our qRT-PCR analysis (Figure 3.1E) showing a decrease in steady-state *Pabpn1* mRNA levels in C2C12 myotubes relative to C2C12 myoblasts (Figure 3.3D). Importantly, these data also demonstrate that the distal PAS is predominantly utilized in both C2C12 myoblasts and C2C12 myotubes (Figure 3.3D), showing that shifts in alternative polyadenylation do not underlie cell type-specific *Pabpn1* regulation for C2C12 myoblasts or myotubes.

3.3.4 The AU-rich element binding protein HuR interacts with the *Pabpn1* 3'UTR *in vitro*

As the long *Pabpn1* 3'UTR is the predominant variant in C2C12 myotubes, we analyzed predicted *cis*-regulatory elements within this 3'UTR. In this analysis we identified multiple putative AU-rich elements (AREs) (Figure 3.3A,E, Supplementary Figure 4). AREs are present in a subset of transcripts, many which are unstable and require precise control of gene expression (20). RNA binding proteins that recognize AREs can stabilize or destabilize the target transcript by blocking or recruiting RNA degradation machinery (20).

The *Pabpn1* 3'UTR contains several predicted ARE motifs including a single Class I ARE (140) present downstream of the first PAS, denoted as ARE1, and three Class III (140,141) U-rich tracts that are located between PAS I and PAS II (Figure 3.3A), which we have designated ARE2, ARE3, and ARE4. As one approach to assess whether these predicted AREs are functional, we examined the genomic sequence conservation of these putative AREs (Figure 3.3E). We determined that these four AREs, including A and T-rich tracts encoding A and U in the mRNA, are highly conserved in *Homo sapiens, Bos taurus*, and *Mus musculus* (Figure 3.3E). Thus, these putative AREs could be important for *Pabpn1* regulation. To further dissect whether these putative AREs function as *cis*-regulatory elements in C2C12 myotubes, we transfected C2C12 myotubes with firefly luciferase reporters encoding the murine *Pabpn1* 3'UTR with a deleted PAS1 (ΔPASI) to ensure that a single polyadenylated product is generated, with each reporter containing a single disrupted ARE. To determine if ARE1 is a functional element in C2C12 myotubes, we generated a mutation where thymines were mutated to

guanosines (142,143), disrupting the ARE. Because AREs 2, 3, and 4 are much longer than ARE1, we generated individual deletions for each ARE. We used these mutated and deleted reporters to determine if disrupting any of these regions in the *Pabpn1* 3'UTR affects expression. Disrupting ARE1 (Supplementary Figure 3B) or ARE2 (Supplementary Figure 3C) does not significantly change luciferase activity relative to control. However, ARE3 (Supplementary Figure 3D) and ARE4 (Supplementary Figure 3E) deletions significantly increase luciferase activity relative to control in C2C12 myotubes, demonstrating that these regions within the *Pabpn1* 3'UTR function as *cis*-regulatory elements that negatively regulate expression.

RNA binding proteins can recognize and bind AREs to regulate mRNA fate. To identify proteins that interact with the *Pabpn1* 3'UTR perhaps through ARE3 or ARE4, we mined available transcriptome-wide cross-linking and immunoprecipitation (CLIP-seq) datasets investigating RNA binding proteins known to bind AREs that are important for regulating muscle biology (144). In these transcriptome-wide studies, the ARE binding protein Human antigen R (HuR) (145) immunoprecipitates the *Pabpn1* transcript in human cervical cancer cells (146) and human embryonic kidney cells (HEK293) (147). These data from HEK293 (147) and HeLa (146) cells identified a single consensus HuR-binding region within the *PABPN1* ARE4 motif (Figure 3.3A,E). While HuR canonically stabilizes target transcripts (148) and positively regulates many transcripts in C2C12 myoblasts (31), HuR can also negatively regulate a subset of transcripts through recruitment of degradation factors to specific target mRNAs (31).

To investigate the potential roles that HuR may play in regulating *Pabpn1* expression in C2C12 myotubes, we tested whether HuR directly interacts with the

Pabpn1 transcript in these cells. We performed crosslinking-immunoprecipitation followed by qRT-PCR (CLIP-qRT-PCR) on UV-crosslinked C2C12 myotubes. As expected, HuR is detected in the immunoprecipitated sample but not in the IgG control (Figure 3.4A). As shown in Figure 3.4B, the *Pabpn1* transcript is enriched with HuR antibody but not with IgG control immunoprecipitation. To confirm HuR interaction with the *Pabpn1* transcript, we also performed RNA immunoprecipitation (RIP) in the absence of crosslinking followed by qRT-PCR. Consistent with results from the CLIP experiment, *Pabpn1* transcript is enriched with HuR immunoprecipitation relative to IgG control in C2C12 myotubes in this experiment (Supplementary Figure 5A,B).

To investigate whether HuR interacts with the *Pabpn1* transcript via the 3'UTR in C2C12 myotubes, we utilized a biotin RNA affinity purification assay. Specifically, we incubated an *in vitro* transcribed biotinylated *Pabpn1* 3'UTR RNA (Supplementary Figure 5C) with C2C12 myotube lysate. After isolating biotinylated RNAs and interacting proteins through binding to streptavidin beads, we immunoblotted for candidate RNA binding proteins including HuR. We analyzed two additional RNA binding proteins, the KH-type splicing regulatory protein (KSRP), which targets ARE-containing transcripts for degradation (149) and co-regulates target transcripts with HuR in C2C12 myoblasts (31) and the CUG triplet repeat RNA binding protein (CUGBP1), an RNA binding protein reported to target AREs (150,151) and immunoprecipitates *Pabpn1* mRNA in C2C12 myoblasts (152). As shown in Figure 3.4C, HuR interacts with the *Pabpn1* 3'UTR but not the control *GAPDH* 3'UTR. However, CUGBP1 and KSRP do not interact with either the *Pabpn1* or *GAPDH* 3'UTR *in vitro* (Figure 3.4C). Overall,

these data demonstrate that HuR directly interacts with the *Pabpn1* transcript in C2C12 myotubes, and that HuR interacts with the *Pabpn1* 3'UTR *in vitro*.

3.3.5 HuR interacts with a specific region within the *Pabpn1* 3'UTR

As previously described, available HuR CLIP-seq datasets (146,147) show an overlapping HuR CLIP site corresponding with ARE4 in the *Pabpn1* 3'UTR (Figure 3.5A), an ARE that acts as a *cis*-regulatory element that negatively regulates expression in C2C12 myotubes. To determine if this site contributes to HuR interaction with the *Pabpn1* 3'UTR in C2C12 myotubes, we exploited the biotin RNA affinity purification assay. We incubated C2C12 myotube lysate with two biotinylated *Pabpn1* 3'UTR RNAs *in vitro* transcribed from firefly luciferase constructs: a control 3'UTR with a deleted PAS I (ΔPAS I) and intact ARE4, and an ARE4 mutant RNA with deleted PAS I and an ARE 4 with thymine-to-guanine mutations disrupting the ARE (142,143) within the HuR CLIP site (Supplementary Figure 6). We then purified biotinylated RNA and interacting proteins using streptavidin beads and immunoblotted for HuR to determine if mutating ARE4 affects HuR interaction with the *Pabpn1* 3'UTR. As shown in Figure 3.5B,C, there is an ~60% reduction in HuR protein interaction with the mutant ARE4 *Pabpn1* 3'UTR relative to the control *Pabpn1* 3'UTR. This significant reduction in HuR interaction demonstrates that ARE4, previously shown to bind HuR in HeLa (146) and HEK293 (147) cells, mediates HuR interaction with the *Pabpn1* 3'UTR in C2C12 myotube lysate *in vitro*.

To further assess whether ARE4, which interacts with HuR in C2C12 myotubes, is a *cis*-regulatory element that regulates *Pabpn1* expression, we conducted luciferase

assays in C2C12 myotubes using a reporter with specific mutations disrupting the HuR CLIP site. We compared relative luciferase activity between two reporters with the *Pabpn1* 3'UTR downstream of firefly luciferase: a control reporter with a deleted PASI, and a mutant reporter with a deleted PASI and mutations disrupting the HuR CLIP site in ARE4. We deleted PASI to ensure that a single polyadenylated product is generated. As shown in Figure 3.5D, luciferase activity increases ~50% with ARE4 mutations relative to control, demonstrating that the region within ARE4 that interacts with HuR negatively regulates expression from the *Pabpn1* 3'UTR in C2C12 myotubes. Overall, these data show that HuR interacts with ARE4 in the *Pabpn1* 3'UTR, and this *cis* element regulates *Pabpn1* expression in C2C12 myotubes.

3.3.6 HuR negatively regulates *Pabpn1* at the RNA and protein levels

To examine the role of HuR in regulating endogenous *Pabpn1* transcript levels, we employed siRNA-mediated depletion of *HuR* in C2C12 myotubes (Figure 3.6A). As previous reports demonstrate that HuR depletion affects C2C12 differentiation (31), we knocked down *HuR* on differentiation day 3 to ensure that early differentiation stages are not affected. Two consecutive days of HuR knockdown cause a ~90% reduction in HuR protein levels (Figure 3.6B). We confirmed that *HuR* knockdown at this stage of C2C12 differentiation does not significantly affect the degree of C2C12 differentiation attained relative to control knockdown by immunoblotting for the late differentiation marker eMyHC (Supplementary Figure 7).

We then tested whether PABPN1 protein levels are affected by *HuR* depletion in C2C12 myotubes. As shown in Figure 3.6C, there is a significant increase in PABPN1

protein levels with *HuR* knockdown relative to control. To determine whether this HuR-mediated PABPN1 regulation occurs at the RNA or protein level, we assessed whether steady-state *Pabpn1* mRNA levels also change with HuR depletion. Steady-state *Pabpn1* mRNA levels significantly increase ~1.5 fold with HuR knockdown (Figure 3.6D). Together these data demonstrate that HuR negatively regulates *Pabpn1* at the RNA and protein levels in C2C12 myotubes.

3.3.7 HuR negatively regulates *Pabpn1 in vivo*

To expand these studies to determine if HuR regulates *Pabpn1* expression *in vivo,* we utilized a muscle-specific *HuR* knockout mouse model ($HuR^{\Delta/\Delta}$) (153,154). In this mouse, loxP sites flank exon 2 of the *HuR* gene and Cre recombinase is expressed under the control of the muscle-specific *MyoD* promoter. Control $HuR^{fl/fl}$ mice are Cre negative and therefore Cre-mediated recombination does not occur and the *HuR* gene is intact. We confirmed that HuR is not detectable in multiple hindlimb muscles including rectus femoris (Figure 3.7A) and gastrocnemius (Supplementary Figure 8) muscles from $HuR^{\Delta/\Delta}$ mice but is abundant in control ($HuR^{fl/fl}$) muscles. Therefore, this HuR knockout mouse model can be utilized to study HuR-mediated regulation of *Pabpn1* expression in muscle *in vivo.*

We isolated primary myoblasts from hindlimb muscles of $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ mice and used qRT-PCR to assess steady-state *Pabpn1* transcript levels. There was no significant difference in steady-state *Pabpn1* transcript levels in $HuR^{\Delta/\Delta}$ myoblasts relative to $HuR^{fl/fl}$ myoblasts (Figure 3.7D). We also immunoblotted for PABPN1 protein

and found no significant difference in PABPN1 protein levels in $HuR^{\Delta/\Delta}$ myoblasts compared to $HuR^{fl/fl}$ myoblasts (Figure 3.7B,C). These data demonstrate that HuR does not regulate steady-state *Pabpn1* transcript and protein levels in primary myoblasts.

Our data (Figure 3.6) suggest that HuR negatively regulates *Pabpn1* in myotubes. To determine whether HuR regulates *Pabpn1* expression in a myotube-specific manner in primary muscle cells, we differentiated primary myoblasts from $HuR^{\Delta/\Delta}$ and $HuR^{fl/fl}$ mice to generate myotubes. As previous reports demonstrate a role for HuR in C2C12 myoblast differentiation (31), we confirmed that these primary $HuR^{\Delta/\Delta}$ myoblasts properly differentiate in the absence of HuR using morphological (Figure 3.7E) and biochemical (Figure 3.7F) analyses. In myotubes, we identified a significant ~1.5 fold increase in *Pabpn1* transcript levels in $HuR^{\Delta/\Delta}$ myotubes relative to $HuR^{fl/fl}$ myotubes (Figure 3.7G). We also immunoblotted for PABPN1 protein and detected a significant ~3-fold increase in PABPN1 protein levels in $HuR^{\Delta/\Delta}$ myotubes relative to $HuR^{fl/fl}$ myotubes (Figure 3.7F,H). Finally, to determine whether HuR could regulate *Pabpn1* expression in muscle tissue, we compared PABPN1 protein levels from $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ rectus femoris muscle. As shown in Figure 3.7I and quantified in Figure 3.7J, PABPN1 protein levels are significantly higher in $HuR^{\Delta/\Delta}$ muscle compared to $HuR^{fl/fl}$ muscle. These results validate our *in vitro* studies and demonstrate that HuR regulates *Pabpn1* expression *in vivo.*

3.3.8 HuR is more cytoplasmic in C2C12 myotubes than C2C12 myoblasts

To begin to dissect the mechanisms underlying HuR-mediated *Pabpn1* regulation in mature muscle, we used immunoblotting to assess whether HuR protein levels change

in C2C12 myoblasts compared to myotubes. There was no significant change in total HuR protein levels in C2C12 myotubes relative to C2C12 myoblasts (Figure 3.8A,B). However, HuR shuttles between the nucleus and cytoplasm (155) raising the possibility that differential localization of HuR could confer myotube-specific regulation of *Pabpn1* expression. To determine whether a change in the steady-state localization of HuR correlates with differential regulation of *Pabpn1* expression in myoblasts and myotubes, we assessed the subcellular localization of HuR in these cells using biochemical fractionation (Figure 3.8C). As expected (156,157), HuR is primarily nuclear in both C2C12 myoblasts and C2C12 myotubes (Figure 3.8C). However, there is an increase in the level of cytoplasmic HuR in C2C12 myotubes compared to C2C12 myoblasts (Figure 3.8C), consistent with previous observations (158-160). As the cytoplasm of myotubes contain high levels of muscle-specific proteins (e.g. embryonic Myosin Heavy Chain, eMyHC, as shown in Figure 3.1C), we assessed the amount of cytoplasmic HuR in C2C12 myotubes relative to another cytoplasmic (161,162), non-muscle-specific protein HSP90 (Figure 3.8C). Using this metric, there is an increase in cytoplasmic HuR in C2C12 myotubes compared to C2C12 myoblasts (Figure 3.8D). These data suggest that a larger cytoplasmic pool of HuR could contribute to HuR-mediated *Pabpn1* regulation specifically in mature muscle.

3.4 Discussion

This study provides insight into how expression of the *Pabpn1* gene is regulated in mature muscle. We identify multiple *cis*-regulatory elements in the *Pabpn1* 3'UTR and

characterize HuR-mediated *Pabpn1* regulation both *in vitro* and *in vivo*. While we identify HuR as one *trans*-acting factor that regulates expression of *Pabpn1*, our data support a model where additional regulatory mechanisms modulate *Pabpn1* expression in mature muscle.

Our previous work showed that PABPN1 levels are low in mature muscle but not in primary myoblasts and non-muscle tissues (58). Using *in vivo* models to uncover the mechanisms underlying this regulation in muscle tissue would be challenging. Although studying *trans*-acting factors *in vivo* is feasible if appropriate mouse models exist, using *in vivo* methods to elucidate the *cis*-regulatory elements that interact with these factors could be challenging and would require production of multiple mouse models. Therefore, we used a mouse cell culture model of mature muscle to identify and define the *cis*-regulatory elements and *trans*-acting factors that modulate *Pabpn1* expression. We validated this *in vitro* model of mature muscle, C2C12 myotubes. The C2C12 myotubes recapitulate all known characteristics of *Pabpn1* gene expression identified *in vivo* (58), including *Pabpn1* transcript instability and low steady-state mRNA and protein levels. Furthermore, our work shows that regulatory mechanisms discovered in this *in vitro* model of mature muscle are conserved *in vivo*. Therefore, this *in vitro* model can be used for future work studying how *cis*- and *trans*-acting factors regulate gene expression in mature muscle.

We defined multiple *cis*-regulatory elements in the *Pabpn1* 3'UTR including two functional AU-rich elements (AREs). AREs are present in only ~8% of mRNAs, and many of these ARE-containing transcripts are short-lived and require precise control of gene expression (20). AREs interact with RNA binding proteins that stabilize or

destabilize the target transcript. We investigated whether several candidate ARE binding proteins that play important roles in muscle biology (144,156) interact with the *Pabpn1* 3'UTR. Our results show that the well-characterized ARE-binding protein HuR binds to the *Pabpn1* 3'UTR. Two other candidate ARE-binding proteins, KSRP and CUGBP1, did not bind the *Pabpn1* 3'UTR *in vitro.* Our analysis does not preclude the possibility that KSRP and/or CUGBP1 could bind elsewhere in the *Pabpn1* transcript, or could regulate *Pabpn1* expression in other cell types. Therefore, we have defined two regulatory elements in the *Pabpn1* 3'UTR and have identified a candidate *trans*-acting factor, HuR, which could regulate the *Pabpn1* transcript in mature muscle.

The ubiquitous ARE-binding protein HuR (145) stabilizes ARE-containing target transcripts by preventing recruitment of RNA degradation machinery (158,159). However, HuR can destabilize specific target transcripts in C2C12 myoblasts (31). Our data demonstrate that ARE3 and ARE4 regulate *Pabpn1* expression in C2C12 myotubes, and that HuR binds specifically to ARE4 *in vitro.* Mutation of ARE4 causes a ~60% decrease in HuR binding to the *Pabpn1* 3'UTR, suggesting that HuR can also bind to other regions within the *Pabpn1* transcript. These data are consistent with previous reports demonstrating that HuR crosslinks to four sites within the *Pabpn1* 3'UTR in HEK293 cells (147). However, only one of these crosslinked regions, corresponding to ARE4, was also identified in HuR crosslinking data from HeLa cells (146) indicating that HuR binding sites within the *Pabpn1* 3'UTR are likely cell-type dependent and may be influenced by additional *trans*-acting factors such as microRNAs or other RNA binding proteins.

This study provides insight into the mechanisms that underlie the muscle-specific decrease in *Pabpn1* transcript and protein levels *in vivo.* We show both *in vitro,* using C2C12 myotubes and primary myotubes, and *in vivo* that HuR directly negatively regulates *Pabpn1* transcript and protein levels. Importantly, experiments comparing primary myoblasts and myotubes show that this HuR-mediated *Pabpn1* regulation is myotube specific, as steady-state *Pabpn1* transcript and protein levels do not change with HuR depletion in myoblasts. These results suggest that differences must exist in the regulation of *Pabpn1* transcript in myotubes compared to myoblasts.

While we detected no change in total HuR levels when we compared myotubes and myoblasts, we did discover a change in HuR localization that could contribute to myotube-specific regulation of *Pabpn1.* Although the steady-state localization of HuR is predominantly nuclear, HuR shuttles between the nucleus and the cytoplasm (155). Phosphorylation can regulate HuR localization (157), which can affect HuR-mediated regulation of RNA target stability as most properly processed mRNAs are degraded in the cytoplasm (163). Our data show that HuR is primarily nuclear in both C2C12 myoblasts and C2C12 myotubes (156). However, we detect more cytoplasmic HuR in C2C12 myotubes compared to myoblasts, which is consistent with previous reports (160) and could contribute to HuR-mediated *Pabpn1* regulation occurring specifically in mature muscle.

Although we explored a role for HuR in regulating *Pabpn1* mRNA levels, additional cytoplasmic functions for HuR include regulation of translation and interplay with microRNAs (164-167). HuR can both promote and repress translation of target transcripts through binding at the 3'UTR (164) including through binding U-rich tracts in

the 3'UTR (165). HuR can also recruit microRNAs to repress translation (166), and reports in C2C12 myoblasts (165) and cancer cells (167) demonstrate that HuR and microRNAs/noncoding RNAs can have antagonistic effects on translation. Furthermore, HuR and microRNA targets commonly overlap in 3'UTRs, suggesting another level of complexity (168). Therefore, the regulatory mechanisms of HuR, including stabilizing/destabilizing transcripts and promoting/repressing translation, are likely influenced by *trans*-acting factors including microRNAs that could interact with the transcript under specific cellular conditions. Additional mechanisms regulating HuR could also be involved, such as HuR cleavage as a mechanism to regulate cytoplasmic HuR accumulation (158). Furthermore, a circular RNA, *CircPABPN1*, generated from *PABPN1* pre-mRNA binds HuR in HeLa cells, perhaps modulating the ability of HuR to bind the *PABPN1* transcript in these cells (169). Overall, the mechanism by which HuR negatively regulates *Pabpn1* expression specifically in mature muscle is likely complex involving multiple RNA binding proteins, microRNAs, and could affect *Pabpn1* regulation at both the RNA- and protein-levels.

Together, this work has exploited multiple models to elucidate mechanisms that confer mature-muscle-specific regulation of the *Pabpn1* transcript. Identification of this regulatory mechanism could provide an opportunity to modulate PABPN1 protein levels perhaps by disrupting HuR recognition of the binding site in ARE4 in the *Pabpn1* 3'UTR. Low PABPN1 protein levels in tissues affected in patients with OPMD could explain why particular muscles are affected in this disease, and exploiting defined regulatory pathways that increase PABPN1 protein levels could be protective for OPMD patients.

3.5 Materials and Methods

Cell culture: Although most experiments utilize the mouse myoblast cell line C2C12 (ATCC CRL-1772), we have also used mouse fibroblasts (NIH/3T3, ATCC CRL-1658), human embryonic kidney cells (HEK293, ATCC CRL-1573) and primary myoblasts harvested from murine hindlimb muscles for particular experiments. Cultured cells were maintained in a humidified incubator with 5% CO_2 at 37°C. Mouse C2C12 myoblasts were cultured in C2C12 growth media (Dulbecco's Modified Eagle's Medium [DMEM] with 4.5 g/L glucose, 10% FBS, 100 U/ml penicillin, 100 U/ml streptomycin). To induce C2C12 differentiation, C2C12 myoblasts were plated on dishes coated with Entactin—Collagen IV—Laminin (ECL; Upstate Biotechnology) in C2C12 differentiation media which was changed every other day. The majority of experiments utilized a 6-day differentiation protocol in which cells were differentiated in DMEM with 4.5 g/L glucose, 1% horse serum, 100 U/ml penicillin, 100 U/ml streptomycin. For one northern blotting experiment (Figure 3.3D), C2C12 myoblasts were differentiated using a 10-day differentiation protocol (DMEM with 1 g/L glucose, 1% horse serum, 100 U/ml penicillin, 100 U/ml streptomycin). Mouse NIH/3T3 fibroblasts and HEK293 cells were cultured in DMEM with 4.5 g/L glucose supplemented with 10% FBS, 100 U/ml penicillin, and 100 U/ml streptomycin. Primary myoblasts were isolated from control ($HuR^{fl/fl}$) and *HuR*-knockout ($HuR^{\Delta/\Delta}$) hindlimb muscles (154) from three male 3-month-old mice per genotype using a protocol yielding cultures >95% myogenic (170). Briefly, muscles were dissected, minced, and digested with 0.1% pronase (EMD Millipore) for 1h at 37°C with mild agitation. The digested material was triturated and filtered through a

100 μm vacuum filter (EMD Millipore). Primary myoblasts were plated on collagen-coated (Bovine Collagen I, Gibco) dishes in primary growth media (Ham's F10, 20% FBS, 5 ng/ml bFGF, 100 U/ml penicillin, 100 U/ml streptomycin). To induce differentiation, primary myoblasts were plated on dishes coated with ECL in primary differentiation media (DMEM with 1 g/L glucose, 1% Insulin—Transferrin—Selenium-A [ITS] supplement [Invitrogen], 100 U/ml penicillin, 100 U/ml streptomycin). Primary myotubes were harvested after 36-48h in differentiation media as indicated.

Immunoblotting: Cells were lysed and muscles homogenized using standard methods (50). After lysis/homogenization, samples were sonicated at 30% output for 10s to shear chromatin. Protein concentrations were determined using a Bradford assay (BioRad), and equal amounts of protein were boiled in reducing sample buffer (250 mM Tris-HCl pH 6.8, 500 mM DTT, 10% SDS, 0.5% Bromophenol Blue, 50% glycerol) and resolved using 4-20% Criterion TGX polyacrylamide gels (BioRad). Proteins were transferred to 0.2 μm nitrocellulose membranes and these membranes were stained with Ponceau S Solution (Sigma-Aldrich) to assess the amount of total protein loaded across samples. Membranes were then blocked in 10% nonfat dry milk in Tris-buffered saline pH 7.4 with 0.1% Tween-20 (TBS-T). Primary antibodies were diluted in 5% nonfat milk in TBS-T and incubated overnight at 4°C. Species-specific horseradish peroxidase conjugated secondary α-IgG antibodies (Jackson ImmunoResearch Laboratories) were used to detect primary antibodies, followed by enhanced chemiluminescence substrate (ECL, Sigma). Chemiluminescence was detected by exposing blots to autoradiography film. The following antibody concentrations were used: α-PABPN1 (1:4000) (50), α-HuR (3A2, 1:10,000) (171), α-eMyHC F1.652 (1:100, Developmental Studies

Hybridoma Bank, Iowa City, Iowa) (172), α-Histone H3 (1:20,000, Abcam), α-KSRP (1:2000, Bethyl), α-CUGBP1 (1:1500, Bethyl), and α-HSP90 (1:5000, Santa Cruz). Ponceau staining and immunoblots were quantified using ImageJ software. To normalize protein loading, Ponceau stain was quantified across a region of the membrane with similar staining patterns for all lanes.

RNA preparation, cDNA and qRT-PCR: Total RNA was isolated using TRIzol reagent (ThermoFisher Scientific) and treated with DNase I, Amplification Grade (Invitrogen) according to the manufacturer's instructions. cDNA was synthesized using the M-MLV reverse transcriptase kit (Invitrogen) and Rnasin (Promega). Approximately 10 ng of cDNA was mixed with appropriate primers and SYBR Select Master Mix (Applied Biosciences) for qRT-PCR analysis. Samples were analyzed using the comparative Ct method (120) on an Applied Biosciences Step One Real Time PCR System. Samples were normalized to *Gapdh* or *Hprt* as indicated.

RNA stability assay: Cells were treated with 5 μg/mL Actinomycin D to inhibit transcription. C2C12 and NIH/3T3 cells were lysed after 30 min, 1h, 2h, 4h, 6h, and 8h following Actinomycin D treatment. RNA was isolated and reverse transcribed as previously described. Transcript levels were determined using qRT-PCR and normalized to levels of *Gapdh* for C2C12, and *Hprt* for 3T3/NIH cells. Fold change was calculated relative to *Pabpn1* levels at the 30 min time point for C2C12 myoblasts and myotubes and 3T3/NIH cells. Each data point representing the mean of 3-4 individual experiments (173) was plotted and appropriately fitted lines/curves corresponded with reasonable correlation coefficients (R^2 >0.9). For the *Pabpn1* transcript in C2C12 myotubes, linear regression was conducted using GraphPad Prism to calculate the slope and the decay rate

constant (k) and mRNA half-life was calculated using the equation $t_{1/2} = \ln(2)/k$ (174). For the *Myc* transcript, RNA half-life was quantified with GraphPad Prism using a one-phase decay equation to fit a nonlinear regression curve to plotted data.

Sequence alignment: Sequences were obtained from the UCSC Genome Browser (175) for the human (Dec. 2013, GRCh38/hg38) (176), bovine (Jun. 2014, Bos_taurus_UMD_3.1.1/bosTau8) (177) and murine (Dec. 2011, GRCm38/mm10) (176) *Pabpn1* 3'UTRs. Sequences were aligned using DNASTAR MegAlign Pro with the MUSCLE alignment algorithm. Uncorrected pairwise distance was used to determine the distance of human, bovine, and murine *Pabpn1* 3'UTRs (Supplementary Figure 3A) with global gap removal and 841 residues considered. Percent identity relative to human *Pabpn1* 3'UTR was calculated as %ID = 100*(1 – distance).

Defining polyadenylation site (PAS) utilization: Primers were designed to amplify regions within the coding DNA sequence (CDS) and the region between the proximal and distal PAS (distal). PAS usage was quantified by calculating the ΔΔCt as distal minus CDS after *Gapdh* normalization (2). Fold change is presented as a ratio relative to the average fold change from C2C12 myoblast samples.

Northern blotting: Total RNA from tissues and cells was isolated as described above under, "RNA preparation, cDNA and qRT-PCR." RNA (20 µg) was loaded on a 1.4% agarose denaturing formaldehyde gel. RNA was transferred to Amersham Hybond-N+ membrane (GE Healthcare and Life Sciences) with 10X SSC buffer. RNA was UV-crosslinked and prehybridized at 65°C in 100 µg/ml salmon sperm DNA (Invitrogen) in Rapid-Hyb buffer (GE Healthcare Life Sciences). PCR using primers spanning exons 2-3 (forward primer) and exons 6-7 (reverse primer) generated an ~400 nucleotide *Pabpn1*

probe (58). This *Pabpn1* PCR product was [α-^{32}P]dCTP (Perkin Elmer) labeled using the Amersham RediPrime II DNA Labelling System (GE Healthcare Life Sciences) and purified with illustra MicroSpin G-25 columns (GE Healthcare Life Sciences). The *Pabpn1* probe was hybridized to the membrane rotating at 65°C overnight. The reverse *18S* oligo was end-labeled with [γ-^{32}P]dATP (Perkin Elmer) using PNK (New England Biolabs) and purified with illustra MicroSpin G-25 columns (GE Healthcare Life Sciences). The *18S* probe was hybridized rotating at 42°C overnight. A Typhoon phosphorimager or autoradiography film were used to detect labeled probe.

Plasmids: The coding sequence for the murine *Pabpn1* 3'UTR was synthesized (GeneArt) and *XhoI/XbaI* digestion was used to insert the fragment into the pcDNA3 luciferase vector (178) by the Emory Integrated Genomics Core (EIGC). Site-directed mutagenesis was used to delete the first polyadenylation site (PASI, 5'-ATTAAA-3'), termed ΔPASI. The ΔPASI 3'UTR construct was used as a template for site-directed mutagenesis to disrupt AREs by mutating thymines to guanosines (142,143). This mutation scheme was used to generate two mutated *Pabpn1* 3'UTR plasmids: one plasmid with mutations disrupting ARE1, and a separate plasmid with mutated ARE4 where thymines that are conserved within the consensus HuR CLIP site (146,147) from mice to humans were mutated (Supplementary Figure 6). The ΔPASI 3'UTR construct was also used to generate three separate plasmids with a deleted ARE2, ARE3, and ARE4, respectively. The pGEM-Teasy *GAPDH* 3'UTR (pGEM-Teasy) plasmid was previously described (179).

Crosslinking and immunoprecipitation of RNA (CLIP) and RNA immunoprecipitation (RIP): To investigate interactions between HuR and target

mRNAs, UV-crosslinking and immunoprecipitation (CLIP) and RNA immunoprecipitation (RIP) were performed on C2C12 myotubes according to standard methods (180,181). Similar procedures were used for both CLIP and RIP, with the addition of a UV-crosslinking step (400 mJ/cm^2, Stratagene UV Stratalinker 2400) before cell lysis in CLIP experiments. Briefly, C2C12 myotubes grown in a 150 cm dish were rinsed twice with ice-cold PBS and lysed with an equal pellet volume of RIPA-2 buffer [50 mM Tris, pH 7.4, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 1% Na-deoxycholate, 1 mM EDTA, and 1 cOmplete protease inhibitor tablet (Sigma)]. Protein-A Dynabeads (Invitrogen) were incubated with either mouse IgG or HuR antibody (171). Beads coated in antibody were resuspended in NT2 buffer (50 mM Tris–HCl, pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% Nonidet P-40 with Rnasin [Promega] and 1 mM DTT). Thawed and clarified lysates were added and the bead/antibody/lysate mixture was incubated at 4°C overnight rotating end-over-end. Beads were washed with cold NT2 buffer five times. Proteinase K treatment released RNAs from bound proteins and input and bound RNA was isolated with TRIzol (Invitrogen) and reverse transcribed as described above. Transcript enrichment was calculated using the comparative Ct method (120), with samples normalized to input and compared to IgG control.

Biotin RNA affinity purification: All *Pabpn1* 3'UTR PCR products generated for use in *in vitro* transcription reactions were amplified from firefly luciferase constructs described in the, "Plasmids" section above. To analyze HuR interaction with the *Pabpn1* 3'UTR (Figure 3.4C), we utilized a construct containing the wildtype murine *Pabpn1* 3'UTR downstream of firefly luciferase. The *Pabpn1* 3'UTR was amplified using PCR with a forward primer containing the T7 promoter. A pGEM-Teasy vector containing the

wildtype human *GAPDH* 3'UTR with a T7 promoter (179) was linearized by *SpeI* digestion (New England Biolabs) and gel purified. For experiments determining the region of HuR interaction with the *Pabpn1* 3'UTR (Figure 3.5B), we used T7 primers to amplify the *Pabpn1* 3'UTR from the following firefly luciferase constructs: ΔPASI murine *Pabpn1* 3'UTR with a deleted PASI, and mutARE4 murine *Pabpn1* 3'UTR with a deleted PASI and mutated ARE4. DNA template (1 μg) was used for *in vitro* transcription. Biotinylated RNAs were generated with the T7 MAXIscript Kit (ThermoFisher Scientific) with biotin-16-UTP (Roche) according to the manufacturer's instructions in the presence of Rnasin (Promega). Biotinylated RNAs were purified with illustra MicroSpin G-25 columns (GE Healthcare Life Sciences). Biotinylated RNA size and integrity was confirmed using RNA gel electrophoresis (Supplementary Figure 5C), which was also used to confirm equal RNA quantification for experiments comparing ΔPASI *Pabpn1* 3'UTR and mutARE4 *Pabpn1* 3'UTR biotinylated RNAs (Supplementary Figure 6). To reduce nonspecific binding, Streptalon streptavidin magnetic beads (bioWORLD) were blocked in 4 mg/ml BSA (Sigma) in IP buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM DTT, 1% Nonidet P-40 supplemented with cOmplete protease inhibitor tablets [Roche Applied Science]) for 3h at 4°C rotating end-over-end. Beads were then incubated with 400 ng biotinylated RNA and Rnasin (Promega) for 20 min at 4°C rotating end-over-end. To prepare lysate, C2C12 myotubes were lysed in IP buffer, sonicated twice for 10s at 30% output, and centrifuged at 21,000 x g for 20 min at 4°C. A total of 400 μg of protein lysate per sample was pre-cleared with washed beads rotating end-over-end at room temperature for 20 min. Lysate, biotinylated RNA, and beads were incubated at room temperature rotating end-over-end

for 20 min. Beads were washed five times with cold IP buffer and proteins were eluted by boiling for 5 min in reducing sample buffer. Input and bound proteins were resolved by SDS-PAGE using 4-20% Criterion TGX polyacrylamide gels (BioRad) and analyzed by immunoblotting.

Firefly luciferase assays: C2C12 myoblasts were plated and differentiated in 96-well plates for luciferase assays. On differentiation day 5, myotubes were transfected with luciferase reporter plasmids using DNA-In CRISPR reagent (Amsbio) according to the manufacturer's instructions in the absence of antibiotics. Media was changed 16h after transfection. Luciferase assays were conducted using the Dual-Luciferase Reporter Assay System (Promega). Cells were lysed 24h after transfection according to the manufacturer's instructions. A BioTek Microplate reader was used to measure luminescence with a 2s delay and 2s integration time. Renilla luciferase (pRL-CMV [Promega]) activity was used to normalize firefly luciferase activity. Data are presented as the ratio of firefly to renilla luciferase activity relative to ΔPASI control for three independent experiments.

siRNA knockdown: Standard siRNA knockdown was employed to deplete HuR. siRNA transfections were conducted using Lipofectamine 2000 (Invitrogen) and siRNAs (HuR Stealth RNAi MSS205314, Low GC negative control [Invitrogen]) were used at a final concentration of 100 nM. C2C12 myotubes were transfected in the absence of antibiotics on the third day of differentiation (differentiation day 3). Media was changed 16h after transfection. To enhance HuR depletion, a second knockdown was conducted 24h after the first transfection (differentiation day 4). Cells were harvested 72h after the initial transfection (differentiation day 6). Immunoblotting was performed to confirm

knockdown, and qRT-PCR was conducted to assess changes in steady-state *Pabpn1* transcript levels with *HuR* knockdown.

Cellular fractionation: C2C12 myoblasts and myotubes were fractionated using standard methods (182). Briefly, cells were scraped in ice-cold PBS and centrifuged at 1000 x g for 5 min. Cell pellets were resuspended in lysis buffer (10 mM Tris, pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% Nonidet P-40, 2 mM vanadyl ribonucleoside complex [VRC]) and incubated on ice for 5 min. Lysate was centrifuged at 1000 x g for 3 min at 4°C to pellet nuclei. The supernatant was collected as cytoplasm and cleared by centrifugation at 13,000 rpm for 10 min at 4°C. The nuclear pellet was washed twice and resuspended in lysis buffer.

Animals: All mice are from the C57Bl/6J background. Muscle-specific HuR knockout ($HuR^{\Delta/\Delta}$) mice were generated by crossing $HuR^{fl/fl}$ mice containing loxP sites flanking exon 2 of the *HuR* gene with mice containing the Cre recombinase gene under the control of the *MyoD* promoter (154). All experiments used muscle tissue from 3-month-old male mice in biological triplicate per genotype. Rectus femoris and gastrocnemius muscles were used for immunoblotting experiments. Hindlimb muscles including rectus femoris, gastrocnemius, soleus, and tibialis anterior muscles were harvested for primary myoblast isolation. Primary myoblasts were isolated from three animals per genotype and treated as independent biological samples. Testing for rodent-related pathogens was routinely performed by McGill University's Comparative Medicine and Animal Resources Center. Animal experiments were approved by the McGill University Faculty of Medicine Animal Care Committee and in accordance with the guidelines set by the Canadian Council of Animal Care.

Data analysis: For the majority of experiments, raw data are normalized to the appropriate values (Ponceau stain for immunoblots, *Gapdh*/*Hprt* for qRT-PCR, renilla luciferase activity for luciferase assays, etc) as indicated. The average normalized values for control samples (C2C12 myoblasts, $HuR^{fl/fl}$cells, etc) was set to 1, and all data are presented as values relative to this mean.

Statistical analysis: For all experiments, statistical analyses were performed using GraphPad Prism. For RNA decay analysis, one-way analysis of variance (ANOVA) was conducted. Student's t-test was used for all other experiments. In all cases, $p < 0.05$ is considered statistically significant.

3.6 Figures

Legends to Figures

Figure 3.1 Steady-state *Pabpn1* mRNA and protein levels are low in C2C12 myotubes relative to C2C12 myoblasts

(**A**, **B**) Representative images of (**A**) C2C12 myoblasts (MB) and (**B**) C2C12 myotubes (MT) are shown. Bar=100μm. (**C**) A representative immunoblot using α-PABPN1 antibody to analyze PABPN1 levels in C2C12 myoblast (MB) and myotube (MT) lysates is shown. The differentiation marker embryonic Myosin Heavy Chain (eMyHC), which is detected only in C2C12 myotube lysate, provides evidence for differentiation. α-Histone 3 (H3) and Ponceau staining serve as loading controls. (**D**) Quantification of PABPN1 protein levels from three independent immunoblots of lysate prepared from C2C12 myoblasts (MB) and C2C12 myotubes (MT) was performed as described in Materials and

Methods. Data were normalized to Ponceau staining as the loading control and are presented as fold change in PABPN1 levels relative to myoblast, which was set to 1.0. Data are mean ± SEM of n=3 samples per cell type (*P<0.05). (**E**) qRT-PCR of *Pabpn1* was used to quantify the steady-state levels of *Pabpn1* transcript in total RNA isolated from C2C12 myoblasts (MB) and C2C12 myotubes (MT). For this experiment, the C2C12 myoblast sample analyzed was then differentiated into the myotube sample that was analyzed. Data are presented as fold change relative to myoblast, which was set to 1.0. For all samples, *Pabpn1* levels were normalized to *Gapdh.* Data presented are mean ± SEM of n=3 samples per cell type (*P<0.05).

Figure 3.2 The *Pabpn1* transcript is unstable specifically in C2C12 myotubes

Cells were treated with Actinomycin D (Act-D) to inhibit transcription and total RNA was collected at the indicated time points post-treatment. To assess transcript stability, qRT-PCR was used to detect *Pabpn1* transcript in these samples. *Myc* stability was determined as a control for an unstable transcript. Levels are presented relative to the amount of transcript present at t=0.5h, which was set to 1.0. For NIH/3T3 cells (**A**), *Pabpn1* and *Myc* levels were normalized to *Hprt*. (**B**) For C2C12 myoblasts (MB) and (**C**) C2C12 myotubes (MT), *Pabpn1* and *Myc* levels were normalized to *Gapdh.* Normalizer transcripts (*Hprt* and *Gapdh*) are set to 1.0 and presented as control stable transcripts on each graph. The calculated half-life of the control unstable *Myc* transcript is ~0.8h in all cell types tested. Because the *Pabpn1* transcript is stable in NIH/3T3 (**A**) and C2C12 myoblasts (**B**) over the course of these experiments, the half-life of the *Pabpn1* transcript in these cells could not be calculated. For C2C12 myotubes (**C**), the half-life of

the *Pabpn1* transcript is ~13h. Data are presented as mean ± SEM of *n*=3-4 independent experiments per cell type. ANOVA was conducted to determine if the mean transcript levels (*Pabpn1*, *Myc*) at each time point are statistically significantly different from the mean transcript level detected at t=0.5h (**P*<0.05).

Figure 3.3 The long *Pabpn1* 3'UTR contains putative conserved *cis*-regulatory elements

(**A**) Schematic of the *Pabpn1* transcript including a 5' untranslated region (5'UTR), the coding DNA sequence (CDS), which encodes the PABPN1 open reading frame, and a 3' untranslated region (3'UTR) which contains two polyadenylation signals (PASI, PASII) and multiple putative AU-rich elements (ARE1, ARE2, ARE3, ARE4). (**B**) Schematic of qRT-PCR strategy using primers that recognize the coding DNA sequence (CDS primers) and distal (distal primers) regions to determine which polyadenylation site (PASI or PASII) is utilized in C2C12 myoblasts (MB) and C2C12 myotubes (MT). (**C**) qRT-PCR was used to quantify *Pabpn1* levels using primers that recognize the *Pabpn1* CDS or the distal *Pabpn1* 3'UTR. The levels of distal 3'UTR-containing transcripts is calculated relative to CDS-containing transcripts and normalized to *Gapdh.* As described in Materials and Methods, these data are presented as fold change relative to C2C12 myoblasts, for which the average value was set to 1.0. Data are mean ± SEM of *n*=3 samples per cell type. (**D**) Representative northern blot using a radiolabeled probe recognizing both short and long forms of the *Pabpn1* transcript, corresponding to PASI and PASII utilization, respectively, assessing PAS usage in multiple mouse tissues and cell lines including C2C12 myoblasts (MB) and C2C12 myotubes (MT) that have been

differentiated for 10 days. Because testis tissue contains high levels of predominantly short *Pabpn1* transcript, this sample was underloaded to avoid strong signal. *18S* serves as a loading control. Asterisks indicate non-specific band. Data are representative of *n*=3 independent experiments. (**E**) Motif analyses of genomic sequences corresponding to putative AREs within the *Pabpn1* 3'UTR to investigate sequence conservation in *Mus musculus, Bos taurus, and Homo sapiens*. Blue box indicates ARE1. Motifs for AREs 2, 3, and 4 show the complete sequence of these putative AREs.

Figure 3.4 HuR binds the *Pabpn1* 3'UTR in C2C12 myotubes

(**A**) Lysates from C2C12 myotubes were immunoprecipitated with α-HuR or control IgG antibody. Immunoblotting was used to detect HuR and the Input sample is also shown. HSP90 is a negative control to show that HuR immunoprecipitation does not purify non-specific proteins. (**B**) HuR-transcript interactions were assessed by crosslinking and RNA immunoprecipitation using qRT-PCR and primers recognizing the *Pabpn1* and *Gapdh* transcripts. Transcript levels are normalized to input and presented as fold change relative to IgG negative control samples. Data are mean ± SEM of *n*=3 samples. (**C**) Biotinylated *GAPDH* and *Pabpn1* RNAs were incubated in C2C12 myotube lysate and the biotinylated RNA along with associated proteins was purified on streptavidin beads. Purified proteins were detected by immunoblotting bound fractions with α-HuR, α-KSRP, and α-CUGBP1 antibodies. The 1% Input lane shows that candidate proteins are present in total C2C12 myotube lysate. Data are representative of *n*=3 independent experiments.

Figure 3.5 HuR binds ARE4, a *cis*-regulatory element that negatively regulates *Pabpn1* expression

(**A**) Schematic illustrating the consensus HuR binding site within the *PABPN1* 3'UTR identified in two transcriptome-wide CLIP-seq studies (146,147). The locations of ARE1, 2, 3, and 4 are indicated as well as PASI and PASII. (**B**) Biotinylated Control (deleted PASI) and Mutant (deleted PASI, mutated HuR CLIP site) *Pabpn1* 3'UTR RNAs were incubated in C2C12 myotube lysate and then purified on streptavidin beads. Immunoblotting with α-HuR and α-KSRP was used to detect purified proteins. The 1% input lane shows that HuR and KSRP are present in total C2C12 myotube lysate. Data represent n=3 independent experiments. (**C**) Quantification of HuR binding to Control and Mutant *Pabpn1* 3'UTR biotinylated RNAs. Data are normalized to 1% Input and fold change was calculated relative to control, which was set to 1.0. Data are presented as mean ± SEM of n=3 independent experiments (*P<0.05). (**D**) C2C12 myotubes were co-transfected with firefly luciferase reporters with control (deleted PASI) and mutant (deleted PASI, mutated HuR CLIP site) *Pabpn1* 3'UTRs and renilla luciferase reporters. Firefly luciferase activity was normalized to renilla luciferase activity as described in Materials and Methods, and the average ratio was set to 1.0 for the control and is presented as mean ± SEM of n=3 independent experiments (*P<0.05).

Figure 3.6 HuR negatively regulates Pabpn1 transcript and protein levels in C2C12 myotubes

C2C12 myotubes were transfected with control siRNA or siRNA targeting the *HuR* transcript. (**A**) Control and *HuR* knockdown C2C12 myotube lysates were

immunoblotted with α-PABPN1 and α-HuR antibodies. Ponceau stain serves as a loading control. The immunoblot shown is representative of *n*=4 independent experiments. (**B**) Quantification of α-HuR immunoblot comparing C2C12 myotube lysate from control and *HuR* knockdown samples. Data are presented as fold change in HuR relative to control knockdown, which was set to 1.0. Data are presented as ± SEM of *n*=4 independent experiments (**P*<0.05). (**C**) Quantification of α-PABPN1 immunoblot from control and *HuR* knockdown samples. Data are presented as fold change in PABPN1 relative to control knockdown, which was set to 1.0. Data are presented ± SEM of *n*=4 independent experiments (**P*<0.05). (**D**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and *HuR* knockdown samples. *Pabpn1* levels are normalized to *Gapdh* and presented as fold change relative to control siRNA samples. Data are presented as ± SEM of *n*=4 independent experiments (**P*<0.05).

Figure 3.7 HuR regulates *Pabpn1* transcript and protein levels specifically in mature muscle *in vitro* and *in vivo*

(**A**) Immunoblot using α-HuR antibody to analyze HuR protein levels in rectus femoris muscle from 3-month-old male control ($HuR^{fl/fl}$) and HuR knockout ($HuR^{\Delta/\Delta}$) mice. For each genotype, samples were prepared in parallel from three independent mice. Ponceau stain serves as a loading control. (**B**) Lysate from primary myoblasts isolated from $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ mice was immunoblotted with α-PABPN1 and α-HuR antibodies. Primary myoblasts were prepared and analyzed in biological triplicate per genotype. Ponceau stain serves as a loading control. (**C**) Quantification of PABPN1 protein levels in $HuR^{fl/fl}$

and $HuR^{\Delta/\Delta}$ primary myoblasts. Data are presented as fold change in PABPN1 levels relative to the mean levels detected in $HuR^{fl/fl}$ myoblasts and plotted as ± SEM of *n*=3 samples. (**D**) qRT-PCR was used to assess steady-state *Pabpn1* transcript levels in $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ primary myoblasts. As described in Materials and Methods, data are normalized to *Gapdh* and are presented as mean fold change relative to mean *Pabpn1* transcript levels detected in $HuR^{fl/fl}$ myoblasts as ± SEM of *n*=3 samples. (**E**) Representative images of primary $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ myotubes. Bar=100μm. (**F**) Lysate from three biological replicates of $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ primary myotubes was immunoblotted with α-PABPN1, α-HuR, and α-eMyHC antibodies. Ponceau stain serves as a loading control. Data are *n*=3 samples. (**G**) qRT-PCR was used to assess steady-state *Pabpn1* transcript levels in $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ primary myotubes. *Pabpn1* transcript levels are normalized to *Gapdh* and fold change is calculated relative to the mean, normalized *Pabpn1* levels observed in $HuR^{fl/fl}$ primary myotubes. Data are presented as mean ± SEM of *n*=3 samples (**P*< 0.05). (**H**) Quantification of immunoblot (**F**) assessing PABPN1 protein levels in $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ primary myotubes. PABPN1 levels are normalized to Ponceau stain and presented as fold change compared to the average levels observed in $HuR^{fl/fl}$ primary myotubes, which was set to 1.0. Data are presented for mean ± SEM of *n*=3 samples (**P*<0.05). (**I**) Lysates from rectus femoris muscle tissue from 3-month-old male $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ mice was immunoblotted with α-PABPN1 to determine PABPN1 protein levels. Results are presented for biological duplicates. Ponceau stain serves as a loading control. (**J**) Quantification of immunoblot analyzing PABPN1 levels in $HuR^{fl/fl}$ and $HuR^{\Delta/\Delta}$ rectus femoris muscle tissue. Data are presented as

fold change relative to the average for the $HuR^{fl/fl}$ sample, which was set to 1.0. Data shown represent the mean ± SEM of n=3 samples (*P<0.05).

Figure 3.8 HuR localization is more cytoplasmic in C2C12 myotubes than C2C12 myoblasts

(**A**) C2C12 myoblast (MB) and C2C12 myotube (MT) lysates were immunoblotted with α-HuR antibody to determine HuR protein levels in these cells. For this experiment the C2C12 myoblast sample analyzed was then differentiated into the myotube sample analyzed for each biological replicate (#1, #2, and #3). Ponceau stain serves as a loading control. (**B**) Quantification of immunoblot shown in (**A**). HuR levels are presented as fold change relative to mean levels detected in C2C12 myoblasts (MB), which was set to 1.0. Data are presented as mean ± SEM of n=3 samples. (**C**) Cellular fractionation was performed on C2C12 myoblasts (MB) and C2C12 myotubes (MT) using three independent biological samples (#1,#2, and #3) for each. Nuclear and cytoplasmic fractions were analyzed by immunoblotting with α-HuR. α-H3 is a control nuclear protein, and α-HSP90 is a control cytoplasmic protein. Ponceau serves as a loading control. (**D**) Quantification of subcellular HuR localization in C2C12 myoblasts (MB) and C2C12 myotubes (MT). Data were normalized to Ponceau and their respective subcellular protein controls (H3 for the nuclear fraction and HSP90 for the cytoplasmic fraction). The ratio of cytoplasmic/total protein detected was calculated and is presented as fold change relative to the mean ratio calculated for C2C12 myoblasts. Data are presented as mean ± SEM of n=3 samples (*P<0.05).

Figure 3.1

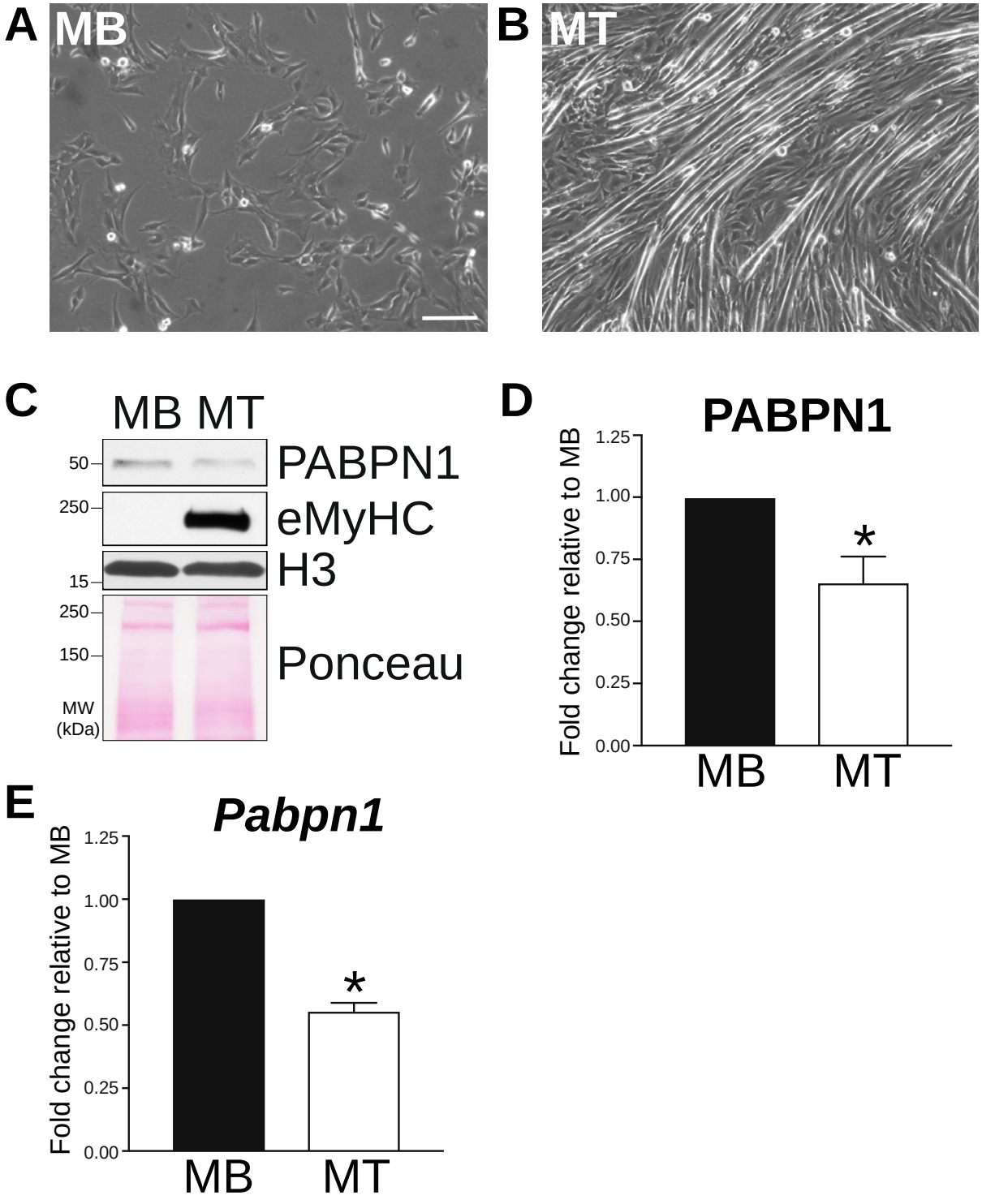

A
MB
B
MT
C
MB MT
50
PABPN1
250
eMyHC
H3
15
250
150
Ponceau
MW
(kDa)
D
PABPN1
Fold change relative to MB
1.25
1.00
0.75
0.50
0.25
0.00
*
MB
MT
E
Pabpn1
Fold change relative to MB
1.25
1.00
0.75
0.50
0.25
0.00
*
MB
MT

Figure 3.2

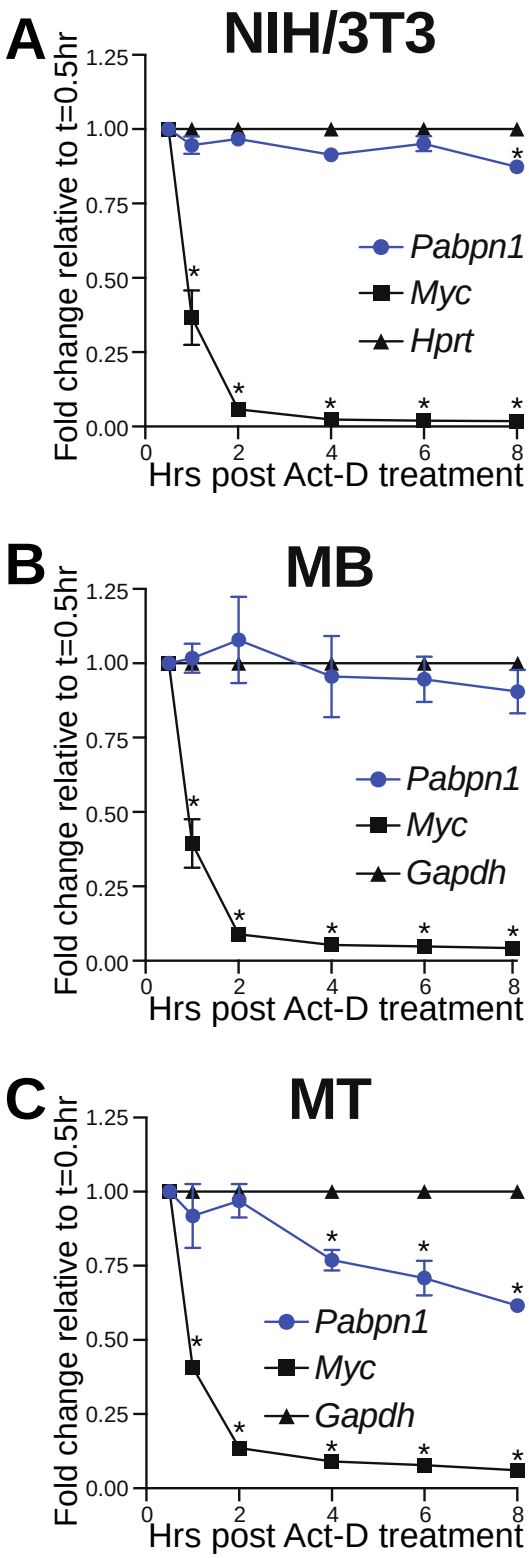

A
NIH/3T3
Fold change relative to t=0.5hr
1.25
1.00
0.75
0.50
0.25
0.00
0
2
4
6
8
Hrs post Act-D treatment
Pabpn1
Myc
Hprt
B
MB
Fold change relative to t=0.5hr
Hrs post Act-D treatment
Pabpn1
Myc
Gapdh
C
MT
Fold change relative to t=0.5hr
Hrs post Act-D treatment
Pabpn1
Myc
Gapdh

Figure 3.3

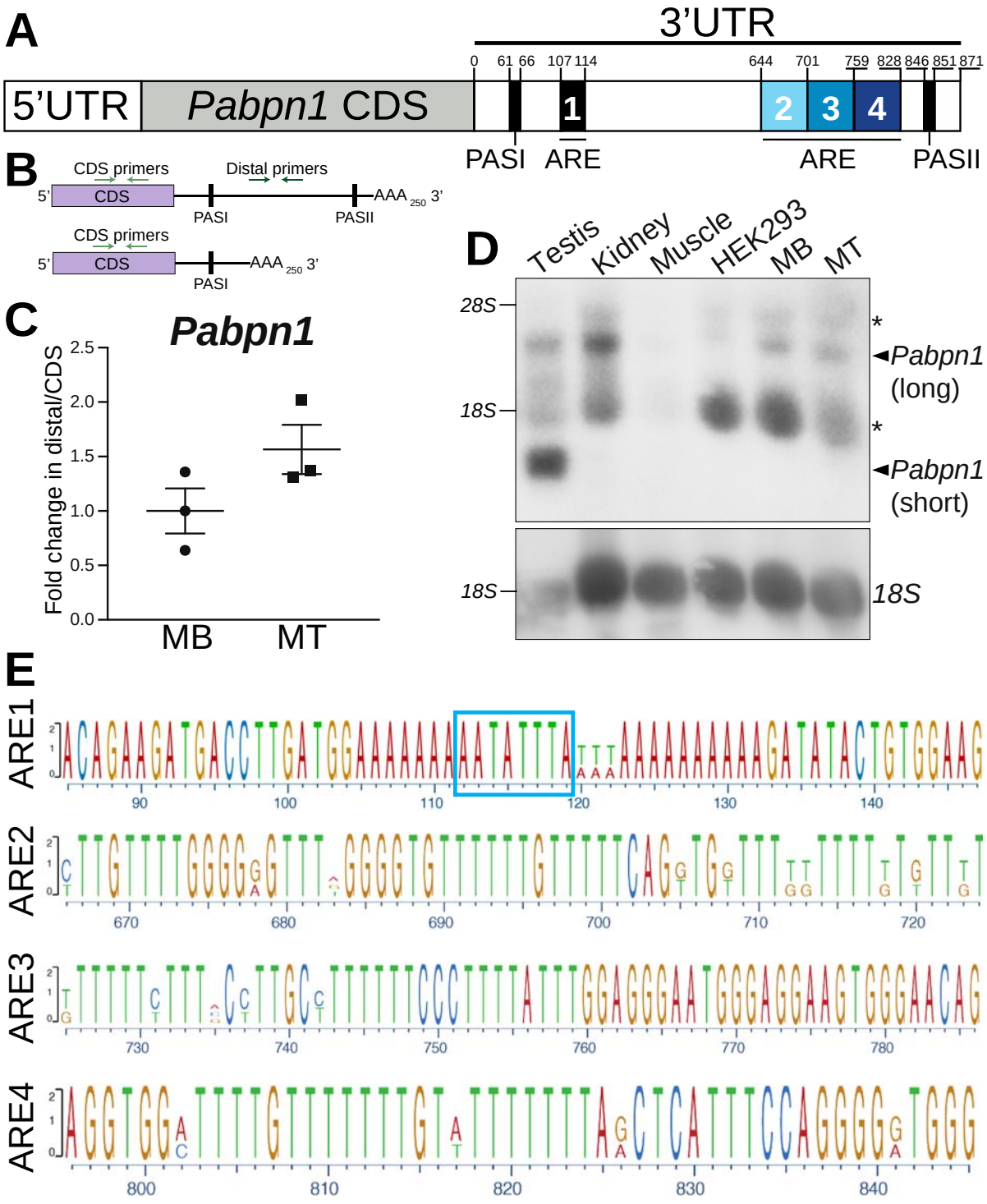

A
3'UTR
0 61 66 107 114 644 701 759 828 846 851 871
5'UTR
Pabpn1 CDS
1
2
3
4
PASI
ARE
ARE
PASII
B
CDS primers
Distal primers
5'
CDS
AAA250 3'
PASI
PASII
CDS primers
5'
CDS
AAA250 3'
PASI
C
Pabpn1
Fold change in distal/CDS
2.5
2.0
1.5
1.0
0.5
0.0
MB
MT
D
Testis
Kidney
Muscle
HEK293
MB
MT
28S
18S
*
◄Pabpn1
(long)
*
◄Pabpn1
(short)
18S
18S
E
ARE1
ARE2
ARE3
ARE4

Figure 3.4

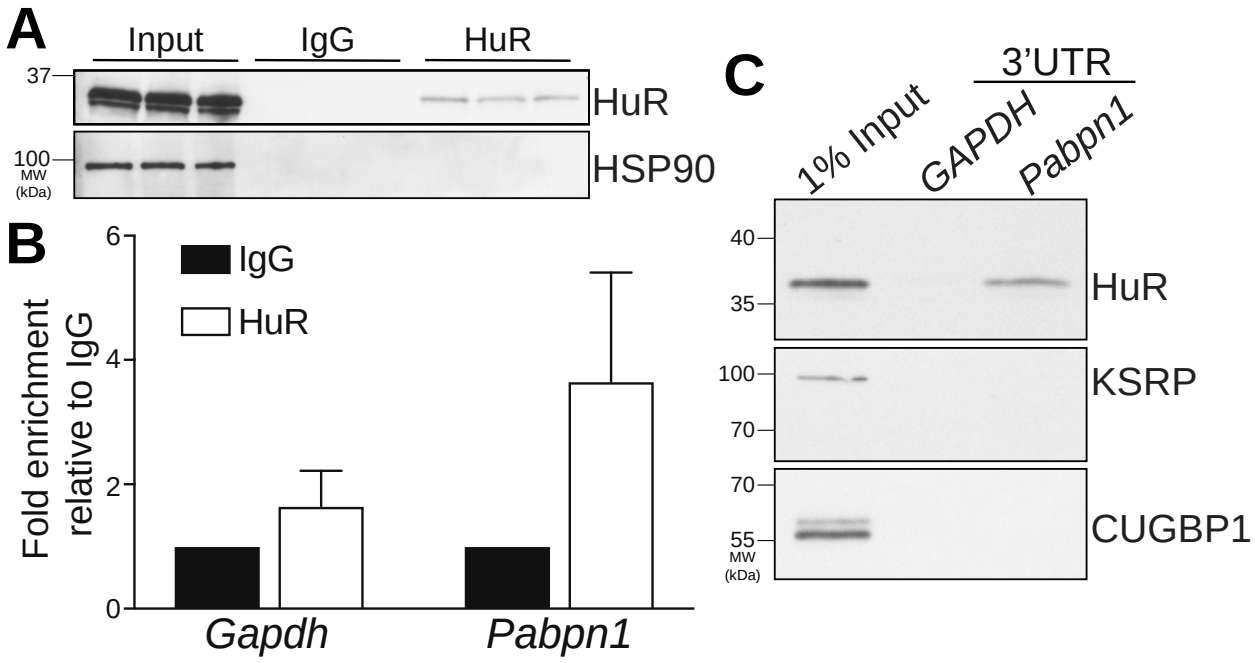

A
Input
IgG
HuR
37
100
MW
(kDa)
HuR
HSP90
B
Fold enrichment
relative to IgG
6
4
2
0
IgG
HuR
Gapdh
Pabpn1
C
1% Input
3'UTR
GAPDH
Pabpn1
40
35
100
70
70
55
MW
(kDa)
HuR
KSRP
CUGBP1

Figure 3.5

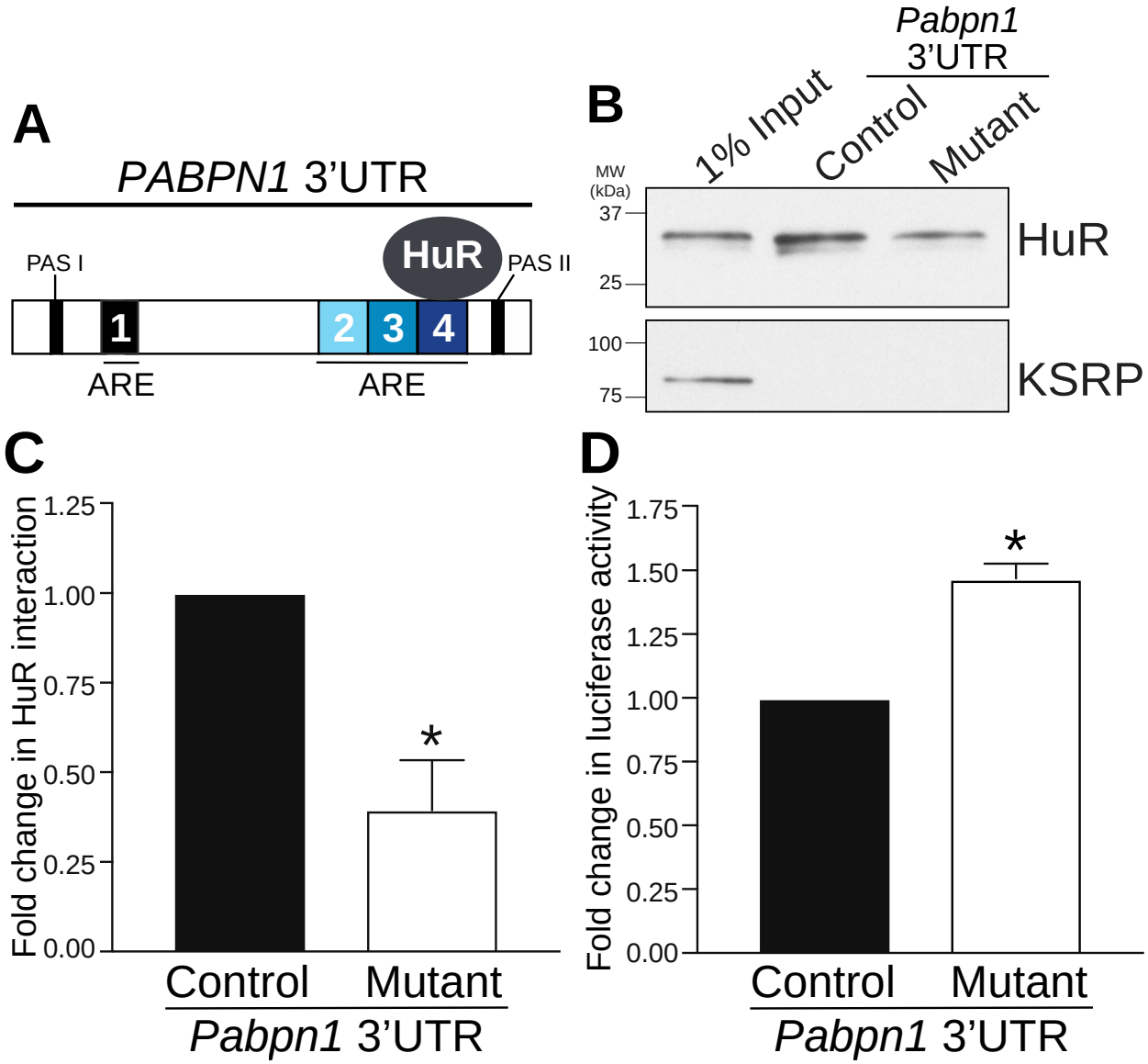

A
PABPN1 3'UTR
PAS I
HuR
PAS II
1
2
3
4
ARE
ARE
B
Pabpn1 3'UTR
1% Input
Control
Mutant
MW (kDa)
37
25
100
75
HuR
KSRP
C
Fold change in HuR interaction
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
Mutant
Pabpn1 3'UTR
D
Fold change in luciferase activity
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
Mutant
Pabpn1 3'UTR

Figure 3.6

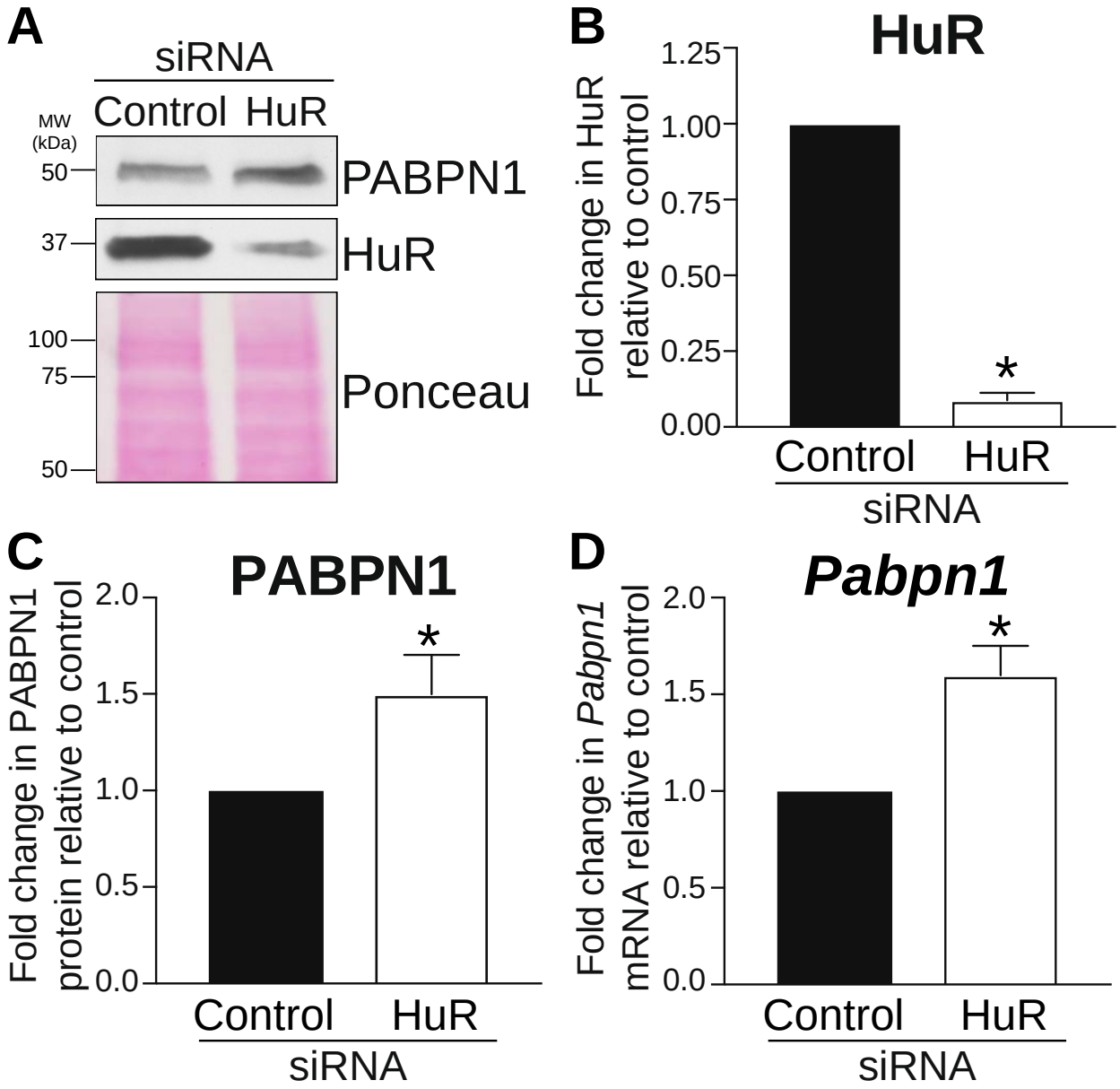

A
siRNA
Control
HuR
MW (kDa)
50
PABPN1
37
HuR
100
75
Ponceau
50
B
HuR
Fold change in HuR relative to control
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
HuR
siRNA
C
PABPN1
Fold change in PABPN1 protein relative to control
2.0
1.5
1.0
0.5
0.0
*
Control
HuR
siRNA
D
Pabpn1
Fold change in Pabpn1 mRNA relative to control
2.0
1.5
1.0
0.5
0.0
*
Control
HuR
siRNA

Figure 3.7

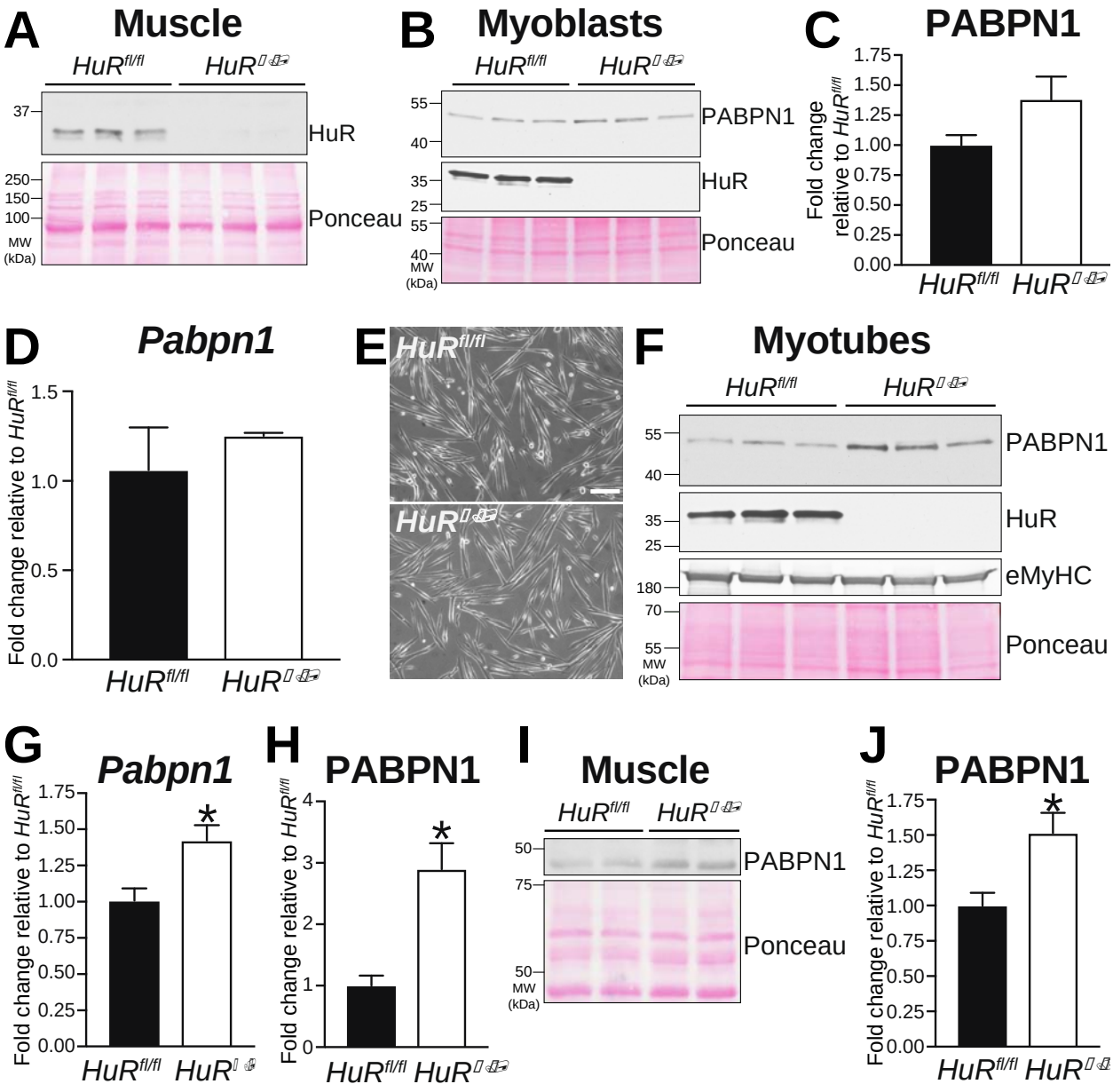
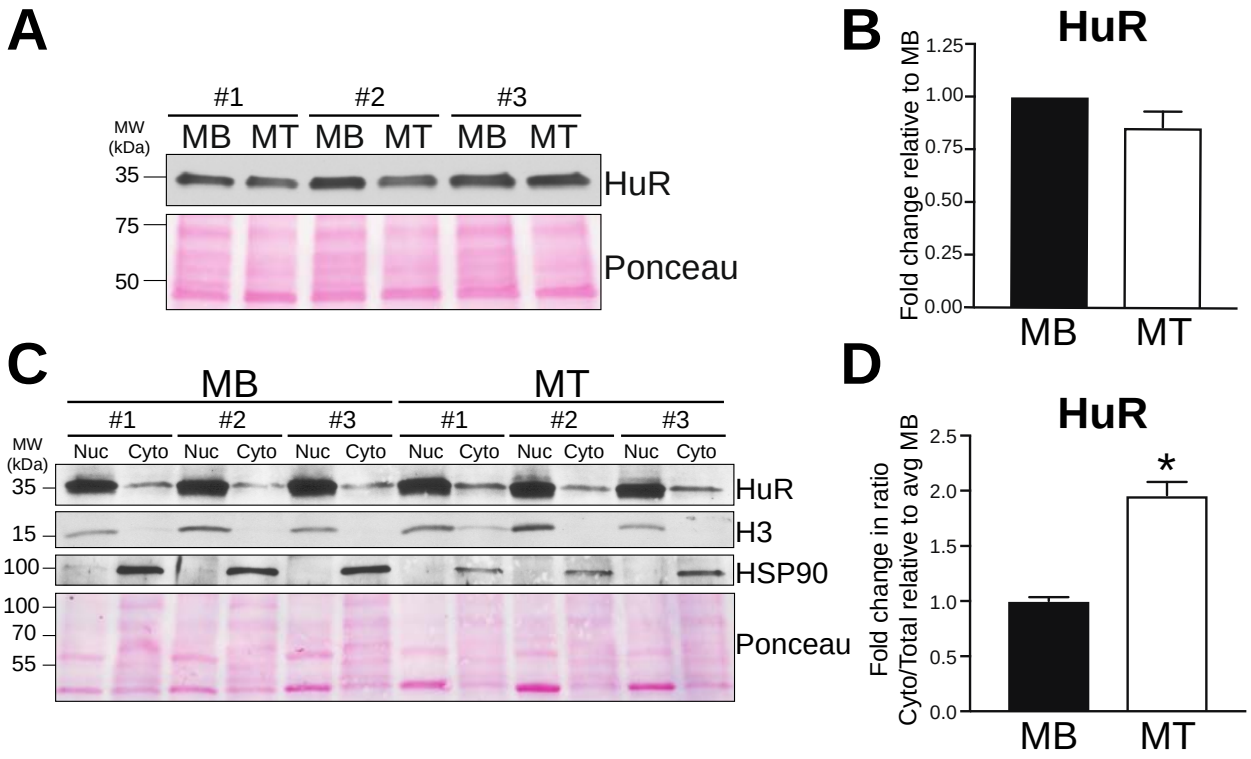

A
Muscle
HuR^fl/fl
HuR^Δ
37
HuR
250
150
100
MW (kDa)
Ponceau
B
Myoblasts
HuR^fl/fl
HuR^Δ
55
40
PABPN1
35
25
HuR
55
40
MW (kDa)
Ponceau
C
PABPN1
Fold change relative to HuR^fl/fl
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
HuR^fl/fl
HuR^Δ
D
Pabpn1
Fold change relative to HuR^fl/fl
1.5
1.0
0.5
0.0
HuR^fl/fl
HuR^Δ
E
HuR^fl/fl
HuR^Δ
F
Myotubes
HuR^fl/fl
HuR^Δ
55
40
PABPN1
35
25
HuR
180
eMyHC
70
55
MW (kDa)
Ponceau
G
Pabpn1
Fold change relative to HuR^fl/fl
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
*
HuR^fl/fl
HuR^Δ
H
PABPN1
Fold change relative to HuR^fl/fl
4
3
2
1
0
*
HuR^fl/fl
HuR^Δ
I
Muscle
HuR^fl/fl
HuR^Δ
50
PABPN1
75
50
MW (kDa)
Ponceau
J
PABPN1
Fold change relative to HuR^fl/fl
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
*
HuR^fl/fl
HuR^Δ

Figure 3.8


A
#1
#2
#3
MW (kDa)
MB MT MB MT MB MT
35
HuR
75
50
Ponceau
B
HuR
Fold change relative to MB
1.25
1.00
0.75
0.50
0.25
0.00
MB
MT
C
MB
MT
#1
#2
#3
#1
#2
#3
MW (kDa)
Nuc Cyto Nuc Cyto Nuc Cyto Nuc Cyto Nuc Cyto Nuc Cyto
35
HuR
15
H3
100
HSP90
100
70
55
Ponceau
D
HuR
Fold change in ratio Cyto/Total relative to avg MB
2.5
2.0
1.5
1.0
0.5
0.0
*
MB
MT

Chapter 4: microRNAs regulate *Pabpn1* expression in muscle cells

Brittany Phillips performed experiments and data analysis for the following figures: Figure 4.1B-E, 4.2A, 4.2C, 4.3A-B, 4.3D, 4.4A, and 4.4C-D. Yue Feng performed data analyses for Figure 4.1A. Hyojung Choo performed experiments and data analyses for Figures 4.2B, 4.3C, and 4.4B.

4.1 Summary

RNA binding proteins are responsible for all RNA processing steps, conferring spatio-temporal control of gene expression. PABPN1 is a ubiquitous RNA binding protein implicated in multiple aspects of RNA processing including 3' end formation and polyadenylation. Small expansion mutations in the ubiquitously-expressed *PABPN1* gene that expand an N-terminal alanine tract from 10 to 11-18 alanines cause the muscle-specific disease oculopharyngeal muscular dystrophy (OPMD). The molecular mechanisms underlying this pathology are still under investigation. PABPN1 protein levels are low in muscle compared to other tissues, which could predispose muscle to OPMD pathology if *PABPN1* mutations further reduce the functional pool of PABPN1 in this tissue. Therefore, modulating PABPN1 levels could be protective against OPMD pathology. Here we identify three microRNAs, miR-141-5p, miR-331-3p, and miR-532-3p, predicted to regulate *Pabpn1* levels in muscle. We use overexpression and depletion approaches to assess whether these microRNAs modulate endogenous *Pabpn1* levels in muscle cells. We determined that miR-141-5p and miR-331-3p negatively regulate steady-state *Pabpn1* transcript levels in muscle cells. Interestingly, miR-532-3p positively regulates *Pabpn1* transcript levels, suggesting that complex interactions between positive and negative regulatory factors modulate *Pabpn1* expression in muscle. Overall, interrogating microRNA pathways that modulate *Pabpn1* expression in muscle cells provides insight into the tissue-specific regulation of PABPN1. Furthermore, these pathways could be exploited to increase PABPN1 protein levels in OPMD.

4.2 Introduction

RNA binding proteins are involved in every aspect of RNA processing, from 5' capping and 3' end formation to eventual turnover in the cytoplasm. Because many of these processing steps are important regardless of cell type, many genes encoding RNA binding proteins are ubiquitously expressed (1). Interestingly, a number of mutations in genes encoding RNA binding proteins cause pathology in nervous and muscle tissue (1,36), suggesting that these tissues have complex gene expression requirements.

The nuclear poly(A) binding protein, PABPN1, is a ubiquitously expressed RNA binding protein (144) that functions in multiple aspects of RNA processing including 3' end formation and polyadenylation (33). PABPN1 is required to achieve proper poly(A) tail length both *in vitro*(45,46,50) and *in vivo* (2), and PABPN1 plays roles in poly(A) site selection (2,41,42,44). Furthermore, *Pabpn1* knockdown in primary muscle cells causes nuclear accumulation of poly(A) RNAs (50), suggesting impaired export of improperly processed transcripts. Thus, the ubiquitously expressed PABPN1 protein has been implicated in many critical aspects of RNA processing.

Despite the general role of PABPN1 in multiple steps of RNA processing, small expansion mutations that expand an N-terminal alanine tract from 10 alanines to 11-18 alanines in the *PABPN1* gene cause the muscle-specific disease oculopharyngeal muscular dystrophy (OPMD), which affects muscles of the eyelids, pharynx, and proximal limbs. Why a subset of skeletal muscles is specifically affected by *PABPN1* mutation while other muscles and tissues are not subject to pathology is not understood. Because insoluble nuclear PABPN1 aggregates are a pathological hallmark of OPMD, it

is possible that sequestration of PABPN1 and other proteins (60,61) could interfere with proper cell function, resulting in pathology(33).

Previous work revealed that steady-state *Pabpn1* transcript and protein levels are low in muscle compared to other tissues with even lower PABPN1 levels detected in muscles affected in OPMD such as pharynx (144). These low PABPN1 levels could predispose muscle to OPMD pathology as mutations could reduce levels of functional PABPN1 in muscle below a critical threshold while other tissues with higher endogenous levels of PABPN1 do not fall below this threshold. These data suggest a loss of function model of OPMD pathology where the level of functional PABPN1 is not sufficient to fulfill critical roles necessary for proper muscle function. Therefore, understanding the mechanisms contributing to low PABPN1 protein levels in muscle is critical as modulating these factors to raise PABPN1 levels could be therapeutic in OPMD.

One mechanism that contributes to low PABPN1 levels in muscle is mRNA stability (144). Post-transcriptional regulation of gene expression often confers fine-tuned spatial and temporal control via *trans*-acting factors that bind *cis*-regulatory elements in the 3' untranslated region (3'UTR) of transcripts. RNA binding proteins and microRNAs that interact with 3'UTRs can influence subcellular localization, translation, and RNA stability, all of which can impact protein levels. Therefore, determining what factors post-transcriptionally regulate the *Pabpn1* transcript is critical both to understand the mechanisms that confer tissue-specific regulation of PABPN1 levels and to identify therapeutic approaches to modulate PABPN1 protein levels in muscle.

Previous work showed that the RNA binding protein HuR negatively regulates *Pabpn1* transcript and protein levels *in vitro* and *in vivo* (183). However, there are likely

multiple RNA binding proteins and microRNAs that regulate *Pabpn1* expression specifically in muscle, and the roles that microRNAs play in regulating *Pabpn1* expression in muscle have not been examined. microRNAs target at least 60% of mammalian transcripts (17) and are involved in numerous developmental and pathological processes (184). These small ~22 nucleotide non-coding RNAs guide the RNA-induced silencing complex (RISC) to target transcripts, and Waston-Crick basepairing between the microRNA seed (positions 2-7) and the target transcript (21) signals for either degradation of the target transcript or translational repression (184) potentially in conjunction with RNA binding proteins. Therefore, microRNAs could contribute to regulation of *Pabpn1* expression in a muscle-specific manner.

Here we present work identifying multiple microRNAs that modulate *Pabpn1* expression in cultured muscle cells. Using a discovery-based bioinformatics approach, we identified three microRNAs, miR141-5p, miR331-3p, and miR532-3p predicted to regulate *Pabpn1* expression in muscle. We find that miR-141-5p and miR-331-3p negatively regulate *Pabpn1* expression in muscle cells, while miR-532-3p positively regulates *Pabpn1* expression. These regulatory pathways are mediated through the *Pabpn1* 3'UTR, suggesting that these microRNAs could directly regulate *Pabpn1* expression. Together, these data suggest that microRNA-mediated *Pabpn1* regulation is complex and provide insight into the multiple pathways that could co-regulate *Pabpn1* expression in muscle.

4.3 Results

4.3.1 The *Pabpn1* 3'UTR contains multiple putative microRNA binding sites that are conserved

Our previous work demonstrates that the *Pabpn1* transcript is unstable in muscle compared to non-muscle tissue (58) contributing to low steady-state *Pabpn1* transcript and protein levels in this tissue. To begin to investigate whether microRNAs are involved in this muscle-specific regulation of the *Pabpn1* transcript, we took a bioinformatics approach to identify putative microRNAs predicted to bind conserved sites within the *Pabpn1* 3'UTR. We have focused our analyses on the *Pabpn1* 3'UTR because microRNAs often target this region of transcripts to post-transcriptionally regulate expression (16). As detailed in Materials and Methods, multiple bioinformatics tools identified three microRNAs, miR-141-5p, miR-331-3p, and miR-532-3p, predicted to regulate the *Pabpn1* transcript via sites within the *Pabpn1* 3'UTR (Figure 4.1A) using the following criteria: presence of microRNA recognition elements that could form microRNA:mRNA heteroduplexes with microRNA seed sequences; conservation of these seed sequences across mouse, rat, and human species; favorable binding energy of these heteroduplexes; and expression of these microRNAs in muscle tissue. Results from these analyses are summarized in Table 4.1. These microRNAs, miR-141-5p and miR-331-3p, each have a single predicted binding site, while miR-532-3p has two predicted binding sites within the *Pabpn1* 3'UTR. All predicted binding sites are present between the two polyadenylation sites (PAS) in the *Pabpn1* 3'UTR (Figure 4.1A). Importantly, previous work demonstrates that the majority of *Pabpn1* transcripts utilize the distal

polyadenylation site (PAS) in muscle (144,183), demonstrating that these predicted binding sites are present in the *Pabpn1* 3'UTR in muscle.

To determine if predicted microRNA seed binding sites within the *Pabpn1* 3'UTR are conserved, we performed sequence alignment comparing mouse, bovine, and human *Pabpn1* 3'UTRs. These microRNA seed sequences are indeed highly conserved (Figure 4.1B-E), suggesting that these elements could contribute to *Pabpn1* regulation. This analysis suggests that these microRNAs and their binding sites within the *Pabpn1* 3'UTR could regulate *Pabpn1* expression in muscle.

4.3.2 microRNA-141-5p negatively regulates *Pabpn1* expression

To investigate whether these microRNAs regulate *Pabpn1* expression in muscle, we exploited a previously validated muscle cell culture model, C2C12 (183). These C2C12 cells proliferate as myoblasts, and serum deprivation causes these myoblasts to differentiate into myocytes and, eventually, the terminally differentiated myotubes. Mature C2C12 myotubes have lower levels of *Pabpn1* transcript than proliferative C2C12 myoblasts (183).

To assess whether miR-141-5p modulates steady-state *Pabpn1* transcript levels, we transfected C2C12 myocytes with miR-141-5p mimics. We observed a significant decrease in steady-state *Pabpn1* transcript levels in miR-141-5p mimic-treated cells compared to control cells (Figure 4.2A). To determine if this effect is specific to muscle cells, we also treated human embryonic kidney cells (HEK293) with miR-141-5p mimics. Preliminary data (n=2) suggest that miR-141-5p overexpression decreases *Pabpn1* transcript levels ~10% in HEK293 cells (Figure 4.2B). However, more work is needed to

determine if these miR-141-5p-mediated effects in C2C12 and HEK293 cells are comparable.

Our bioinformatics work (Figure 4.1) suggests that miR-141-5p binds a motif within the *Pabpn1* 3'UTR. To investigate whether these miR-141-5p-mediated effects on *Pabpn1* transcript levels in C2C12 cells are mediated through the *Pabpn1* 3'UTR, we transfected C2C12 cells with a luciferase reporter containing with the *Pabpn1* 3'UTR present downstream of firefly luciferase. We observed a significant ~25% decrease in luciferase activity when C2C12 cells were treated with miR-141-5p mimics compared to control (Figure 4.2C), demonstrating that this effect on *Pabpn1* expression is mediated through the *Pabpn1* 3'UTR. Together these data show that miR-141-5p negatively regulates *Pabpn1* transcript levels, and that this regulation is mediated through the *Pabpn1* 3'UTR.

4.3.3 microRNA-532-3p positively regulates *Pabpn1* expression

Our bioinformatics analyses predicted two miR-532-3p binding sites within the *Pabpn1* 3'UTR (Figure 4.1A). To determine whether overexpression of this miR-532-3p affects endogenous *Pabpn1* transcript levels, we transfected C2C12 myocytes with miR-532-3p mimics. We observed no significant change in steady-state *Pabpn1* transcript levels with miR-532-3p overexpression relative to control-transfected cells (Figure 4.3A). However, treating C2C12 myotubes with miR-532-3p antagomiRs, which deplete endogenous miR-532-3p levels, significantly decreases endogenous *Pabpn1* transcript levels ~15% relative to control-treated samples (Figure 4.3B). Preliminary results (n=2) suggest that this miR-532-3p-mediated regulation is specific to C2C12 cells as miR-532-

3p inhibition in HEK293 cells does not change steady-state *Pabpn1* transcript levels (Figure 4.3C). To determine if this effect on *Pabpn1* transcript levels is mediated through the *Pabpn1* 3'UTR, we transfected C2C12 myocytes with a firefly luciferase reporter containing the *Pabpn1* 3'UTR. We detected a significant ~25% decrease in luciferase activity in C2C12 myocytes when miR-532-3p is inhibited (Figure 4.3D), demonstrating that this regulation is dependent on the *Pabpn1* 3'UTR. Overall, these results suggest that miR-532-3p positively regulates *Pabpn1* transcript levels via the *Pabpn1* 3'UTR.

4.3.4 microRNA-331-3P negatively regulates *Pabpn1* expression

To assess whether miR-331-3p also regulates *Pabpn1* expression, we treated C2C12 myocytes with miR-331-3p mimics. We identified a significant ~30% decrease in endogenous *Pabpn1* transcript levels with miR-331-3p overexpression relative to control treated cells (Figure 4.4A). Interestingly, preliminary data suggest that miR-331-3p overexpression in HEK293 cells only decreases steady-state *Pabpn1* transcript levels ~10% (Figure 4.4B), suggesting that miR-331-3p could elicit stronger effects on *Pabpn1* expression in muscle cells than non-muscle cells. To further dissect the role that miR-331-3p plays in modulating *Pabpn1* expression, we treated C2C12 myotubes with miR-331-3p antagomiRs to reduce the levels of endogenous miR-331-3p. Consistent with miR-331-3p mimic studies, we observed a significant increase in steady-state *Pabpn1* transcript levels in response to miR-331-3p inhibition. To determine if this negative regulation occurs via the *Pabpn1* 3'UTR, we transfected C2C12 cells with firefly luciferase reporters containing the *Pabpn1* 3'UTR. We detected a ~70% decrease in luciferase activity with miR-331-mimic treatment relative to control (Figure 4D). Overall,

these data demonstrate that miR-331-3p negatively regulates *Pabpn1* expression via the *Pabpn1* 3'UTR.

4.4 Discussion

The work presented here investigates microRNA-mediated regulation of *Pabpn1* expression in muscle cells. We determined that miR-141-5p and miR-331-3p negatively regulate endogenous *Pabpn1* transcript levels in muscle cells, and that these effects are mediated through the *Pabpn1* 3'UTR. Interestingly, miR-532-3p positively regulates *Pabpn1* transcript levels in muscle cells via the *Pabpn1* 3'UTR. Future work is needed to determine the exact mechanisms underlying these regulatory pathways, and whether they are potential targets for OPMD therapies to raise PABPN1 levels.

Our bioinformatics analyses identified miR-141-5p as a putative regulator of *Pabpn1* expression in muscle. Although no studies have investigated the roles of miR-141-5p in muscle, previous work has linked miR-141-5p overexpression to various cancers (185,186). To determine if miR-141-5p regulates *Pabpn1* expression in muscle, we transfected C2C12 myocytes with miR-141-5p mimics. We observed a significant decrease in endogenous *Pabpn1* transcript levels in miR-141-5p mimic treated cells compared to control treated cells.

Previous work determined that muscle-specific post-transcriptional regulation of *Pabpn1* contributes to low PABPN1 protein levels (144,183). To determine if this miR-141-5p-mediated effect on *Pabpn1* expression is specific to C2C12 muscle cells, we treated HEK293 cells with miR-141-5p mimics. Preliminary data suggest that this miR-141-5p mimic treatment decreases steady-state *Pabpn1* transcript levels in HEK293 cells.

However, further investigation is needed to determine if these results are statistically significant, and whether miR-141-5p regulates *Pabpn1* expression in a muscle-specific manner.

miR-532-3p has also been predicted to target the *Pabpn1* transcript. Similar to miR-141-5p studies, few molecular studies have investigated miR-532-3p with most studies focusing on miR-532-3p as a biomarker (187). To determine if miR-532-3p could regulate *Pabpn1* expression, we treated C2C12 myocytes with miR-532-3p mimics. We observed no significant change in steady-state *Pabpn1* transcript levels with miR-532-3p mimic treatment relative to control-treated myocytes. However, we detected a significant decrease in *Pabpn1* expression with miR-532-3p depletion with antagomiRs relative to control, indicating that miR-532-3p positively regulates *Pabpn1* expression in C2C12 myotubes. However, this regulation of *Pabpn1* transcript levels is not observed when HEK293 cells are miR-532-3p depleted, suggesting a complex regulatory network that is likely cell-type dependent. Because microRNA mimic experiments were conducted in C2C12 myocytes, which are less differentiated than the C2C12 myotubes used in antagomiR experiments, these data could indicate a differentiation-dependent mechanism of miR-532-3p-mediated *Pabpn1* regulation. Alternatively, if miR-532-3p binding sites in the *Pabpn1* 3'UTR are saturated in C2C12 myocytes, overexpressing this microRNA could have negligible effects on *Pabpn1* expression while miR-532-3p depletion could have a relatively strong effect. Finally, we used reporter assays to begin to dissect whether this miR-532-3p-mediated *Pabpn1* regulation is direct or indirect. We determined that this regulation is mediated through the *Pabpn1* 3'UTR, demonstrating that the 3'UTR is sufficient for miR-532-3p-mediated *Pabpn1* regulation. However, more

work is required to further interrogate the mechanisms underlying miR-532-3p-mediated *Pabpn1* regulation in C2C12 cells and whether miR532-3p regulates PABPN1 protein levels perhaps by modulating translation.

Similar to miR-141-5p and miR-532-3p, miR-331-3p has been identified as a biomarker and/or prognostic indicator for various diseases including prostate cancer (188), glioblastoma (189), hepatocellular carcinoma (190), and laryngeal squamous cell carcinoma (191). However, little is known about the molecular roles of miR-331-3p in muscle. Our work has determined that miR-331-3p mimic-treated C2C12 myocytes have significantly decreased steady-state *Pabpn1* transcript levels compared to control-treated cells. Preliminary data investigating whether this regulatory mechanism is detected in non-muscle cells suggest that miR-331-3p has a larger impact on *Pabpn1* expression in muscle cells than HEK293 cells. However, further investigation is needed to determine if these effects in HEK293 cells are significant. Results from antagomiR experiments conducted in C2C12 myotubes are consistent with data from miR-331-3p mimic experiments and indicate that miR-331-3p negatively regulates *Pabpn1* expression. Together, these data demonstrate that miR-331-3p negatively regulates *Pabpn1* expression via the *Pabpn1* 3'UTR.

Overall, these results are consistent with a model where a complement of *trans*-acting factors including microRNAs and RNA binding proteins coalesce to modulate *Pabpn1* expression in muscle cells. The work presented here has investigated whether predicted microRNAs regulate steady-state *Pabpn1* transcript levels. However, because microRNAs can also modulate translation, future work is needed to dissect how these microRNAs regulate PABPN1 protein levels to determine whether modulation of

translation could also be involved. Although these factors could individually negatively or positively regulate *Pabpn1* expression, the overall *Pabpn1* expression profile is generated as a result of the cumulative effects of these factors. Our work indicates that all microRNAs assessed here modulate *Pabpn1* expression through the *Pabpn1* 3'UTR. Future studies are needed to determine whether these effects are direct or indirect as these microRNAs could modulate factors that in turn regulate the *Pabpn1* transcript. Our work also begins to probe whether these regulatory pathways are dynamic across cell types, which is critical for understanding the mechanisms that regulate *Pabpn1* specifically in muscle and contribute to low PABPN1 protein levels in this tissue. These investigations are critical as increasing PABPN1 protein levels in OPMD could be protective against pathogenesis.

4.5 Materials and methods

microRNA prediction. The DIANA-microT-CDS version 5.0 (http://www.microrna.gr/microT-CDS/) algorithm was used to identify putative microRNA recognition elements in the *Mus musculus Pabpn1* 3'UTR based on binding predictions and element conservation. A high precision threshold of 0.6 was used to identify microRNAs with high miTG prediction scores, with higher scores indicating a higher probability of targeting. We then used RNA22 (https://cm.jefferson.edu/rna22v1.0/) to further filter predictions generated from DIANA-microT-CDS. Briefly, RNA22 uses a pattern-based algorithm to discover microRNA target sites and corresponding heteroduplexes without relying on cross-species

conservation. TargetScan (http://www.targetscan.org) was then used to identify microRNAs with conserved seed sequences across *Homo sapiens, Mus musculus, and Rattus norvegicus* species. RNAHybrid (192) then calculated the folding energy of microRNA:mRNA heteroduplexes. microRNA expression profiles from muscle tissue (193,194) were then assessed to determine whether these microRNAs are expressed in muscle. Results from these analyses are summarized in Table 1.

Sequence alignment. Sequences were obtained from the UCSC Genome Browser (175) for the human (Dec. 2013, GRCh38/hg38) (176), bovine (Jun. 2014, Bos_taurus_UMD_3.1.1/bosTau8) (177) and murine (Dec. 2011, GRCm38/mm10) (176) *Pabpn1* 3'UTRs. Sequences were aligned using DNASTAR MegAlign Pro with the MUSCLE alignment algorithm.

Cell culture. Cultured cells were maintained in a humidified incubator with 5% CO_2 at 37°C. Mouse C2C12 myoblasts were cultured in C2C12 growth media (Dulbecco's Modified Eagle's Medium [DMEM] with 4.5 g/L glucose, 10% FBS, 100 U/ml penicillin, 100 U/ml streptomycin). To induce C2C12 differentiation for microRNA mimic and antagomiR transfections, C2C12 myoblasts were plated on dishes coated with Entactin—Collagen IV—Laminin (ECL; Upstate Biotechnology) in C2C12 differentiation media (DMEM with 4.5 g/L glucose, 1% horse serum, 100 U/ml penicillin, 100 U/ml streptomycin). For luciferase assays, C2C12 myotubes were differentiated as previously described but in the absence of ECL-coating. HEK293 cells were cultured in DMEM with 4.5 g/L glucose supplemented with 10% FBS, 100 U/ml penicillin, and 100 U/ml streptomycin.

RNA preparation, cDNA and qRT-PCR. Total RNA was isolated using TRIzol reagent (ThermoFisher Scientific) and treated with DNase I, Amplification Grade (Invitrogen) according to the manufacturer's instructions. cDNA was synthesized using the M-MLV reverse transcriptase kit (Invitrogen) and Rnasin (Promega). Approximately 10 ng of cDNA was mixed with appropriate primers and SYBR Select Master Mix (Applied Biosciences) for qRT-PCR analysis. Samples were analyzed using the comparative Ct method (120) on an Applied Biosciences Step One Real Time PCR System. Samples were normalized to 18S (mimic and antagomiR in C2C12) or *Gapdh* (mimic and antagomiR in HEK293) as indicated.

microRNA mimics, antagomiRs, and plasmids. For microRNA mimics experiments, C2C12 myoblasts were plated in differentiation media on Day 0. Day 1 myocytes were transfected with mirVana microRNA mimics at a final concentration of 20 nM (mmu-miR-141-5p [MC12692], hsa-miR-331-3p [MC10881], hsa-miR-532-3p [MC12824] mirVana miRNA mimic Negative Control #1 [4464058], Ambion) using Lipofectamine2000 (Invitrogen) according to the manufacturer's instructions in the absence of antibiotics. Myocytes were harvested on Day 2 in TRIzol. For antagomiR experiments, C2C12 myoblasts were plated in differentiation media on Day 0. Day 5 myotubes were transfected with mirVana microRNA mimics at a final concentration of 30 nM (hsa-miR-331-3p [MH10881], hsa-miR-532-3p [MH12824], mirVana miRNA inhibitor Negative Control #1 [4464076], Ambion) using Lipofectamine2000 (Invitrogen) according to the manufacturer's instructions in the absence of antibiotics. Myotubes were harvested on Day 6 in TRIzol. The coding sequence for the murine *Pabpn1* 3'UTR was synthesized (GeneArt) and *XhoI/XbaI* digestion was used to insert the fragment into the

pcDNA3 luciferase vector (178) by the Emory Integrated Genomics Core (EIGC) as previously described (183).

Luciferase assays. C2C12 myoblasts were plated for differentiation (day 0) as previously described. C2C12 myocytes (day 1) were co-transfected with luciferase reporters using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions in the absence of antibiotics. Cells were harvested 24hr after transfection. Luciferase assays were conducted using the Dual-Luciferase Reporter Assay System (Promega). Cells were lysed 24h after transfection (day 2) according to the manufacturer's instructions. A TD-20/20 luminometer (Turner Designs) was used to measure luminescence with a 2s delay and 2s integration time. Renilla luciferase (pRL-CMV [Promega]) activity was used to normalize firefly luciferase activity. Data are presented as the ratio of firefly to renilla luciferase activity relative to control transfected samples for three independent experiments.

Data analysis. For the majority of experiments, raw data are normalized to the appropriate values (*18S* or *Gapdh* for qRT-PCR, renilla luciferase activity for luciferase assays, etc) as indicated. The average normalized values for control samples was set to 1, and all data are presented as values relative to this mean.

Statistical analysis. For all experiments, statistical analyses were performed using GraphPad Prism. Student's t-test was used for all experiments. In all cases, $p < 0.05$ is considered statistically significant.

4.6 Figures

Legends to Figures

Figure 4.1: The *Pabpn1* 3'UTR contains multiple conserved putative microRNA binding sites. (**A**) Sequence of the murine *Pabpn1* 3'UTR. The *Pabpn1* 3'UTR contains multiple putative *cis*-regulatory elements. Two polyadenylation sites (PASI, PASII) determine where cleavage and polyadenylation occur. Multiple putative microRNA binding sites are present between these two PAS sites. (**B-E**) Motif analyses of genomic sequences corresponding to putative microRNA binding sites within the *Pabpn1* 3'UTR to investigate sequence conservation in *Mus musculus, Bos taurus, and Homo sapiens*. Motifs for miR-141-5p, miR-331-3p, and both sites for miR-532-3p show the complete sequence recognized by the seed of these microRNAs.

Figure 4.2: microRNA-141-5p overexpression negatively regulates *Pabpn1* expression. (**A**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-141-5p mimic treated C2C12 samples. *Pabpn1* levels are normalized to 18S and presented as fold change relative to control treated samples. Data are presented as ± SEM of n=4 independent experiments (*P<0.05). (**B**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-141-5p mimic treated HEK293 samples. *Pabpn1* levels are normalized to *Gapdh* and presented as fold change relative to control treated samples. Data are presented as ± SEM of n=2 independent experiments. (**C**) C2C12 cells were co-transfected with firefly luciferase reporters with wildtype *Pabpn1* 3'UTR, renilla luciferase reporters, and miR-141-5p mimics or control. Firefly luciferase activity was normalized to renilla luciferase

activity as described in Materials and Methods, and the average ratio was set to 1.0 for the control and is presented as mean ± SEM of *n*=3 independent experiments (**P*<0.05).

Figure 4.3: microRNA-532-3p positively regulates *Pabpn1* expression. (**A**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-532-3p mimic treated C2C12 samples. *Pabpn1* levels are normalized to 18S and presented as fold change relative to control treated samples. Data are presented as ± SEM of *n*=4 independent experiments. (**B**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-532-3p antagomiR treated C2C12 samples. *Pabpn1* levels are normalized to 18S and presented as fold change relative to control treated samples. Data are presented as ± SEM of *n*=3 independent experiments (**P*<0.05). (**C**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-532-3p antagomiR treated HEK293 samples. *Pabpn1* levels are normalized to *Gapdh* and presented as fold change relative to control treated samples. Data are presented as ± SEM of *n*=2 independent experiments. (**D**) C2C12 cells were co-transfected with firefly luciferase reporters with wildtype *Pabpn1* 3'UTR, renilla luciferase reporters, and miR-532-3p antagomiRs or control. Firefly luciferase activity was normalized to renilla luciferase activity as described in Materials and Methods, and the average ratio was set to 1.0 for the control and is presented as mean ± SEM of *n*=3 independent experiments (**P*<0.05).

Figure 4.4: microRNA-331-3p negatively regulates *Pabpn1* expression. (**A**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-

331-3p mimic treated C2C12 samples. *Pabpn1* levels are normalized to 18S and presented as fold change relative to control treated samples. Data are presented as ± SEM of n=4 independent experiments (*P<0.05). (**B**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-331-3p mimic treated HEK293 samples. *Pabpn1* levels are normalized to *Gapdh* and presented as fold change relative to control treated samples. Data are presented as ± SEM of n=2 independent experiments. (**C**) qRT-PCR was used to assess the steady-state levels of the *Pabpn1* transcript from control and miR-331-3p antagomiR treated C2C12 samples. *Pabpn1* levels are normalized to 18S and presented as fold change relative to control treated samples. Data are presented as ± SEM of n=3 independent experiments (*P<0.05). (**D**) C2C12 cells were co-transfected with firefly luciferase reporters with wildtype *Pabpn1* 3'UTR, renilla luciferase reporters, and miR-331-3p mimics or control. Firefly luciferase activity was normalized to renilla luciferase activity as described in Materials and Methods, and the average ratio was set to 1.0 for the control and is presented as mean ± SEM of n=3 independent experiments (*P<0.05).

Table 4.1: Summary of bioinformatics predictions. Table 4.1 summarizes the results of multiple bioinformatics analyses described in 4.5 Materials and Methods. HS= *Homo sapiens*, RN= *Rattus norvegicus*, MM= *Mus musculus*, Y= predicted by the program, N= not predicted by the program.

Figure 4.1

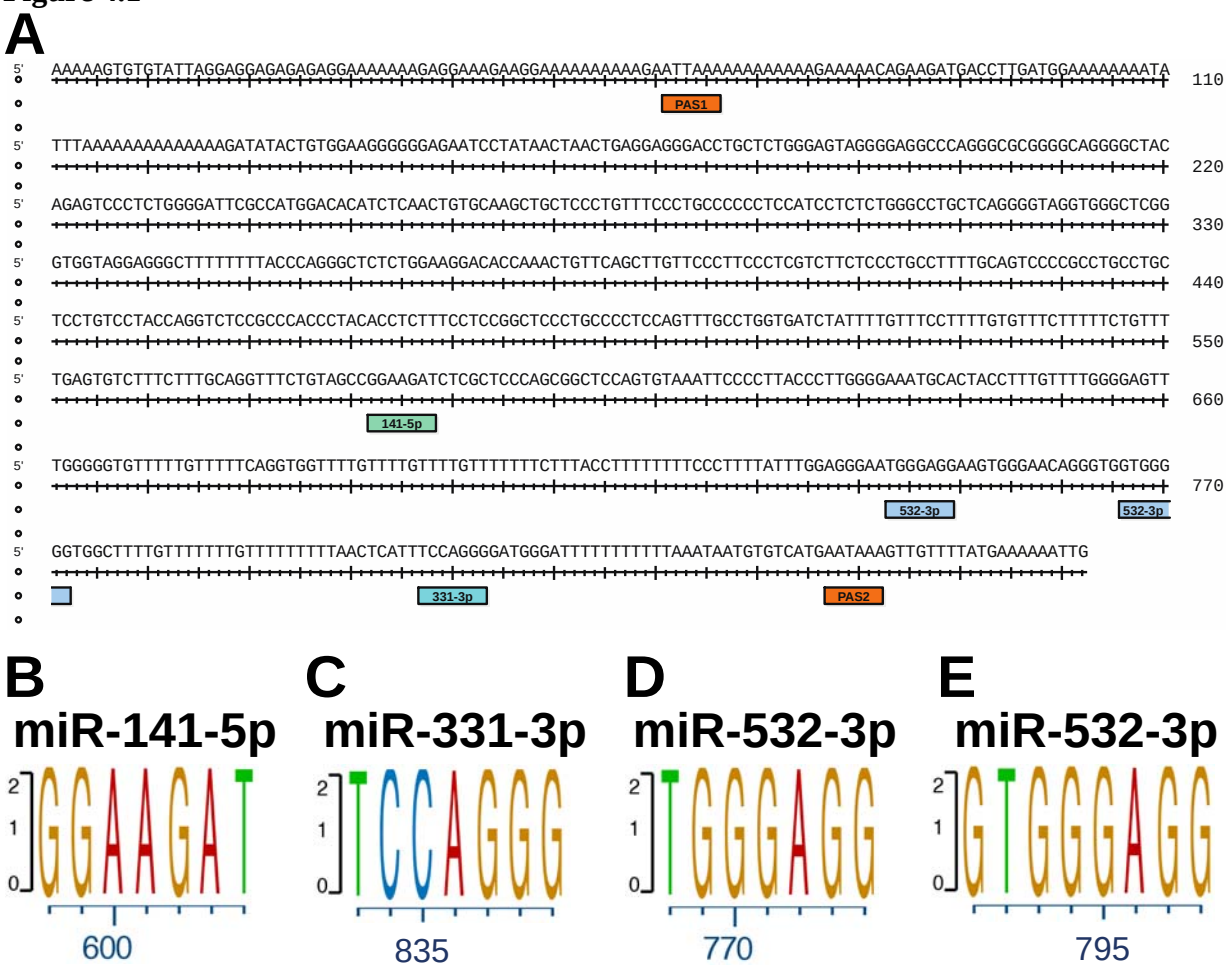

A
5' AAAAAGTGTGTATTAGGAGGAGAGAGAGGAAAAAAAGAGGAAAGAAGGAAAAAAAAAAAGAATTAAAAAAAAAAAAAGAAAAACAGAAGATGACCTTGATGGAAAAAAAATA 110
PAS1
5' TTTAAAAAAAAAAAAAAAGATATACTGTGGAAGGGGGGAGAATCCTATAACTAACTGAGGAGGGACCTGCTCTGGGAGTAGGGGAGGCCCAGGGCGCGGGGCAGGGGCTAC 220
5' AGAGTCCCTCTGGGGATTCGCCATGGACACATCTCAACTGTGCAAGCTGCTCCCTGTTTCCCTGCCCCCCTCCATCCTCTCTGGGCCTGCTCAGGGGTAGGTGGGCTCGG 330
5' GTGGTAGGAGGGCTTTTTTTTACCCAGGGCTCTCTGGAAGGACACCAAACTGTTCAGCTTGTTCCCTTCCCTCGTCTTCTCCCTGCCTTTTGCAGTCCCCGCCTGCCTGC 440
5' TCCTGTCCTACCAGGTCTCCGCCCACCCTACACCTCTTTCCTCCGGCTCCCTGCCCCTCCAGTTTGCCTGGTGATCTATTTTGTTTCCTTTTGTGTTTCTTTTTCTGTTT 550
5' TGAGTGTCTTTCTTTGCAGGTTTCTGTAGCCGGAAGATCTCGCTCCCAGCGGCTCCAGTGTAAATTCCCCTTACCCTTGGGGAAATGCACTACCTTTGTTTTGGGGAGTT 660
141-5p
5' TGGGGGTGTTTTTGTTTTTCAGGTGGTTTTGTTTTGTTTTGTTTTTTTCTTTACCTTTTTTTTCCCTTTTATTTGGAGGGAATGGGAGGAAGTGGGAACAGGGTGGTGGG 770
532-3p
532-3p
5' GGTGGCTTTTGTTTTTTTGTTTTTTTTTAACTCATTTCCAGGGGATGGGATTTTTTTTTTTAAATAATGTGTCATGAATAAAGTTGTTTTATGAAAAAATTG
331-3p
PAS2
B
miR-141-5p
C
miR-331-3p
D
miR-532-3p
E
miR-532-3p
600
835
770
795

Figure 4.2

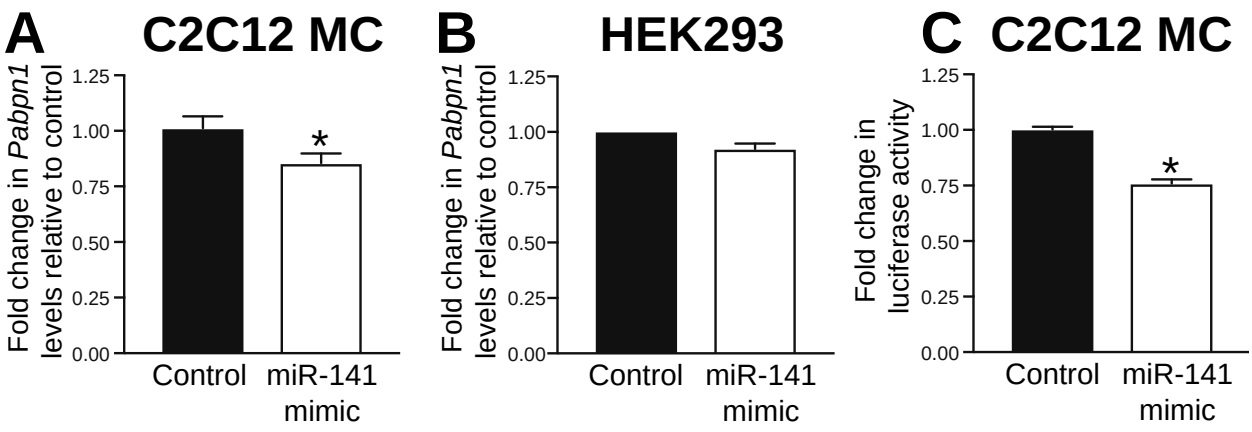

A
C2C12 MC
Fold change in Pabpn1 levels relative to control
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
miR-141 mimic
B
HEK293
Fold change in Pabpn1 levels relative to control
1.25
1.00
0.75
0.50
0.25
0.00
Control
miR-141 mimic
C
C2C12 MC
Fold change in luciferase activity
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
miR-141 mimic

Figure 4.3

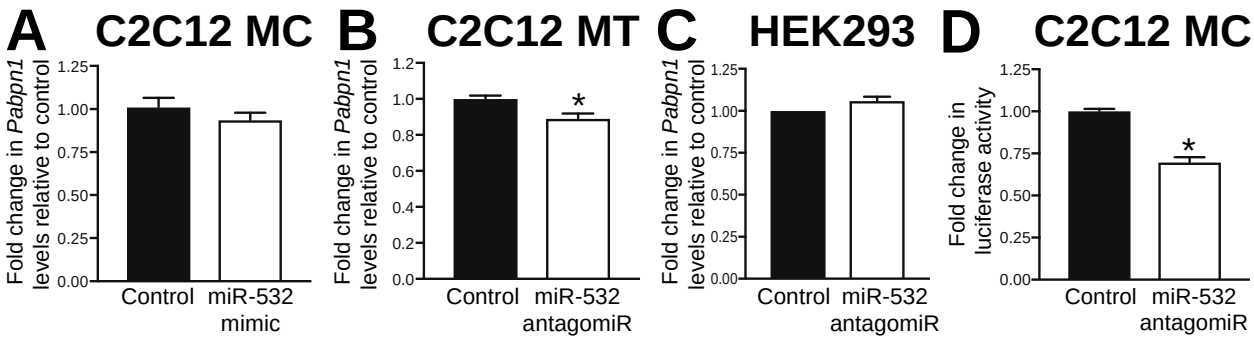

A
C2C12 MC
Fold change in Pabpn1 levels relative to control
1.25
1.00
0.75
0.50
0.25
0.00
Control
miR-532 mimic
B
C2C12 MT
Fold change in Pabpn1 levels relative to control
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
Control
miR-532 antagomiR
C
HEK293
Fold change in Pabpn1 levels relative to control
1.25
1.00
0.75
0.50
0.25
0.00
Control
miR-532 antagomiR
D
C2C12 MC
Fold change in luciferase activity
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
miR-532 antagomiR

Figure 4.4

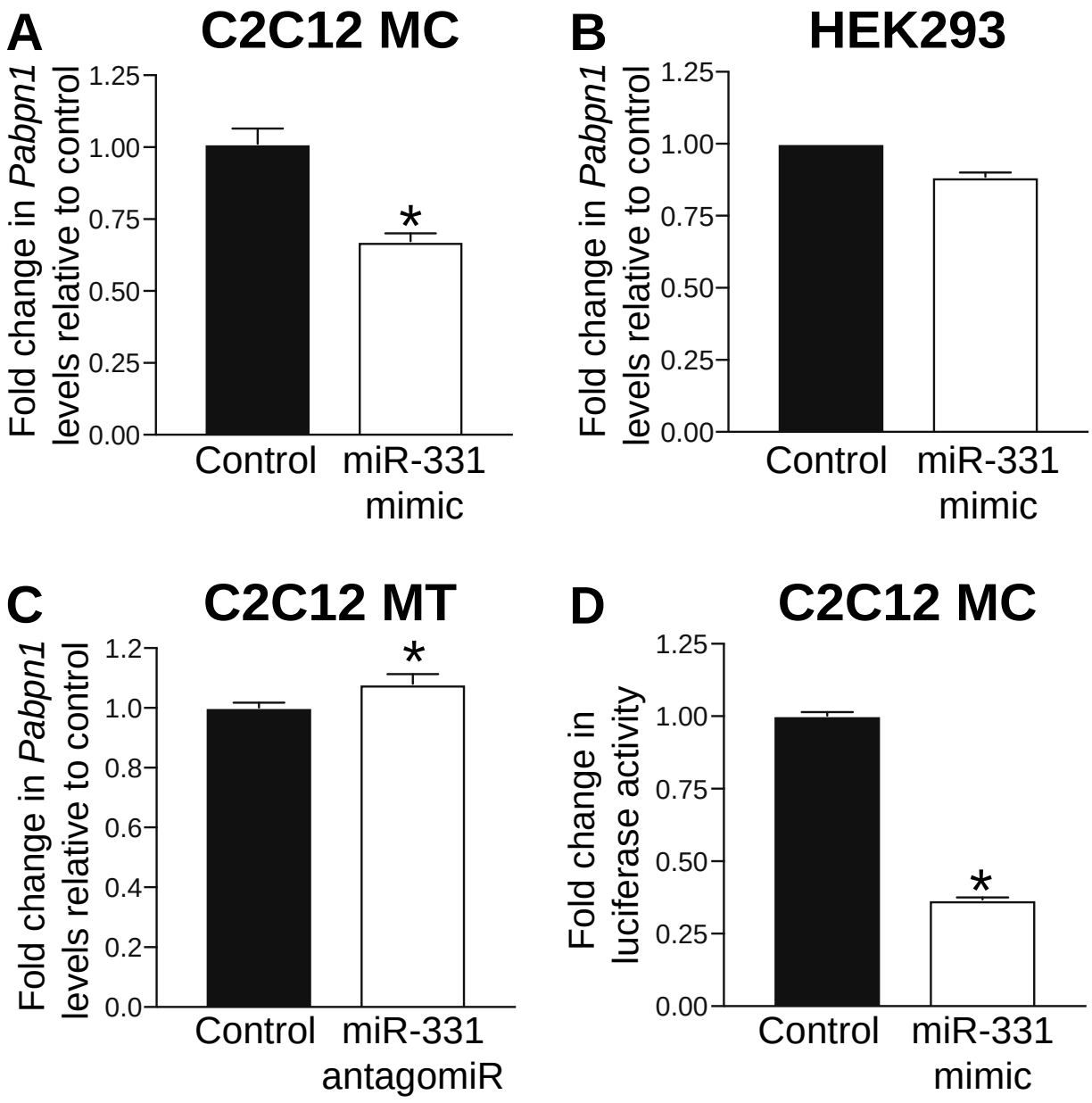

A
C2C12 MC
Fold change in Pabpn1 levels relative to control
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
miR-331 mimic
B
HEK293
Fold change in Pabpn1 levels relative to control
1.25
1.00
0.75
0.50
0.25
0.00
Control
miR-331 mimic
C
C2C12 MT
Fold change in Pabpn1 levels relative to control
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
Control
miR-331 antagomiR
D
C2C12 MC
Fold change in luciferase activity
1.25
1.00
0.75
0.50
0.25
0.00
*
Control
miR-331 mimic

Table 4.1

	DIANA-microT	RNA22	Conservation HS/RN/MM	Folding energy (Kcal/mol)	Muscle expression
miR-141-5p	Y	N	Y	-21	High
miR-331-3p	Y	Y	Y	-20.9	Low
miR-532-5p	Y	Y	Y	-28.3, -24.6	High

Chapter 5: Discussion

5.1 Overview

Regulation of gene expression is critical for proper function at the molecular, cellular, and tissue levels. Although some ubiquitously expressed genes are expressed at similar levels across cells and tissues, gene expression requirements can differ depending on the cellular context. Post-transcriptional regulation is a critical component conferring this fine-tuned cell or tissue-specific expression, often in a spatio-temporal manner. Various *trans*-acting factors including RNA binding proteins and microRNAs mediate this post-transcriptional regulation by modulating RNA processing events such as localization, stability, and translation which in turn influence protein levels. Unsurprisingly, as many of these RNA processing events are important regardless of cell or tissue type, many genes encoding RNA binding proteins are ubiquitously expressed. However, despite these ubiquitous expression profiles, mutations in these genes often cause tissue-specific diseases (1). Understanding the dynamic functions of RNA binding proteins in across tissues could provide insight into why tissues are susceptible to particular disease pathologies.

The work presented in this dissertation seeks to explore models that explain the tissue-specific pathology observed in OPMD. The findings lay the groundwork for understanding how endogenous levels of expanded PABPN1 interfere with proper muscle function. We also investigate mechanisms that regulate *Pabpn1* expression that confer low PABPN1 protein levels specifically in muscle, which could predispose this tissue to pathology. Future studies will utilize these novel OPMD mouse models to study additional processes that could be disrupted in OPMD. These mouse models and primary

muscle cells cultured from these models also allow us to mechanistically investigate the functions and dysfunctions of the alanine-expanded PABPN1 protein. Furthermore, we can utilize these mouse models to design and test therapeutic strategies *in vivo* to determine whether modulating PABPN1 levels in muscle is protective against OPMD phenotypes. Overall, our work emphasizes the importance of studying RNA binding proteins within the context of particular cells or tissues as previously unappreciated tissue-specific roles could be relevant to human diseases.

5.2 OPMD studies: The limitations of overexpression models

It has been twenty years since the link between small expansion mutations in *PABPN1* and OPMD pathology was discovered (38). Since this discovery, several OPMD models have been generated in order to understand how these mutations in the *PABPN1* gene that expand an N-terminal alanine tract in the PABPN1 protein from 10 alanines to 11-18 alanines cause muscle-specific disease (56). Many studies have focused on investigating insoluble nuclear PABPN1 aggregates as the presence of these aggregates in a subset of patient muscle nuclei is a pathological hallmark of OPMD (33). These studies often utilize transgenic overexpression of expanded *PABPN1* in cell culture systems or mice, with the most commonly used mouse model of OPMD overexpressing a transgenic expanded *Pabpn1* allele ~10-30 fold (65).

While this mouse model is a powerful tool for screening drugs that could modulate aggregate formation in muscle nuclei (56), using these models for molecular studies is problematic for several reasons. Previous work demonstrates that steady-state

PABPN1 protein levels are low in muscle relative to other tissues (144), suggesting that these low levels could predispose muscle to pathology when PABPN1 is perturbed. Therefore, overexpressing expanded PABPN1 could confound results as extremely high levels of PABPN1 protein in muscle is not physiologically relevant. Many studies do not control for this overexpression by comparing overexpressed expanded PABPN1 to a mouse overexpressing wildtype PABPN1. Therefore, phenotypes due to PABPN1 overexpression and phenotypes due to PABPN1 expansion cannot be discerned. Novel mouse models are needed to begin to dissect physiologically relevant phenotypes associated with endogenous expression of expanded *Pabpn1*.

Using these overexpression models to understand aggregate dynamics and how these aggregates contribute to OPMD pathology also has limitations. Aggregate-mediated diseases would logically cause pathology in tissues with high levels of mutant protein. These high protein levels could increase protein-protein interactions necessary for seeding and propagating aggregates. Therefore, overexpressing expanded PABPN1 could confound studies as aggregate dynamics could change with high protein levels. Furthermore, because PABPN1 is an RNA binding protein present at extremely low levels in muscle (144), the molecular roles for PABPN1 are likely tightly regulated in this tissue. Therefore, overexpression of mutant PABPN1 could also confound molecular experiments. Overall, while these models of OPMD have generated critical findings for understanding phenotypes related to overexpression of expanded PABPN1 *in vivo*, it is important to understand the specific roles for PABPN1 in muscle under physiologically-relevant conditions.

5.3 Understanding how loss versus gain of PABPN1 function relates to OPMD

PABPN1 mutations only affect a subset of skeletal muscles, suggesting that the requirements for PABPN1 are critical for particular muscles, while other tissues, including other skeletal muscles, are able to properly function when *PABPN1* is mutated. The molecular deficiencies resulting from endogenous expression of expanded PABPN1 in muscle cells and tissue have not been examined except for a few patient muscle samples, which are often collected from unaffected muscle. Therefore, studying both wildtype and expanded PABPN1 at endogenous levels in the context of muscle is necessary to understand why this tissue is specifically affected when PABPN1 is mutated.

Chapter 2 describes the generation and characterization of two novel mouse models of OPMD. The first model contains a ubiquitous heterozygous knock-in of expanded *Pabpn1* allele at the native locus and one wildtype allele (*Pabpn1*$^{+/A17}$). This model is a critical tool for the field as it is the closest-available genocopy to OPMD patients which allows us to determine whether endogenous levels of expanded PABPN1 cause pathology in muscle tissue *in vivo.* We also generated a second conditional *Pabpn1* knockout mouse with one *Pabpn1* allele which, after Cre recombination, has exons 1 and 2 excised, yielding a ~50% reduction in PABPN1 protein levels in all tissues tested (*Pabpn1*$^{+/\Delta}$, Chapter 2). By comparing phenotypes between *Pabpn1*$^{+/A17}$ animals and *Pabpn1*$^{+/\Delta}$ animals, we are able to assess the extent to which expanded PABPN1 is functional, and if particular dysfunctions are likely to contribute to OPMD pathology. Although some phenotypes diverge between *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ animals, we also

observe convergence of phenotypes including poly(A) tail defects, suggesting that expanded PABPN1 is only partially functional, and that loss of function could contribute to OPMD pathology. Low PABPN1 levels in muscle could predispose this tissue to OPMD pathology as mutations interfering with PABPN1 function or sequestration of PABPN1 in aggregates could reduce PABPN1 molecules that are able to fulfill critical roles in muscle. Therefore, understanding the mechanisms that contribute to low PABPN1 levels specifically in muscle is critical. Chapters 3 and 4 describe multiple mechanisms contributing to post-transcriptional *Pabpn1* regulation. In Chapter 3, we determined that the RNA binding protein HuR binds a specific motif in the *Pabpn1* 3'UTR and negatively regulates *Pabpn1* transcript and protein levels *in vitro* and *in vivo* specifically in mature muscle. This study is important because modulating *Pabpn1* transcript and protein levels by targeting this HuR-mediated regulation could be exploited as a novel therapeutic approach to elevate PABPN1 levels in OPMD muscle, which could be protective against pathology. Chapter 4 identifies microRNAs that regulate endogenous *Pabpn1* transcript levels in immortalized muscle cells. We determined that three microRNAs, miR-141-5p, miR-331-3p, and miR-532-3p regulate *Pabpn1* expression via the 3'UTR. Together, this work suggests that the complement of *trans*-acting factors including RNA binding proteins and microRNAs interact with the *Pabpn1* 3'UTR to coordinate tissue and cell-type specific regulation of *Pabpn1* expression (Figure 5.1). This work is significant because it is the first to interrogate specific mechanisms that contribute to low PABPN1 protein levels in muscle, which could predispose this tissue to OPMD pathology. Designing pharmacologic interventions that modulate these regulatory pathways could raise PABPN1 protein levels in muscle, which

could be a therapeutic approach for OPMD treatment. Overall, the work presented in this dissertation begin to dissect how tissue-specific requirements underlie pathology as an approach to better understand molecular dysfunctions in human disease.

5.4 Future directions

Data presented in Chapter 2 show that the *Pabpn1* gene is essential as homozygous *Pabpn1*$^{\Delta/\Delta}$ animals are not generated. Interestingly, homozygous *Pabpn1*$^{A17/A17}$ animals are not produced in Mendelian ratios, suggesting that expanded PABPN1 is not able to fulfill the basic functions of PABPN1 required to support viability and that PABPN1 could play critical roles in development. Together, these data suggest that one copy of wildtype *Pabpn1* is necessary for viability. However, we crossed *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice in an attempt to generate *Pabpn1*$^{A17/\Delta}$ mice which have one expanded *Pabpn1* allele and one *Pabpn1* allele that is knocked out. Surprisingly, preliminary data suggest that one copy of expanded *Pabpn1* is sufficient to rescue these viability defects (data not shown). These preliminary data suggest that expanded *Pabpn1* is partially functional, consistent with our findings in Chapter 2 as only a subset of phenotypes found in *Pabpn1*$^{+/A17}$ mice are recapitulated in *Pabpn1*$^{+/\Delta}$ mice. These data also suggest that dosage of expanded *Pabpn1* is important, as mice with a single expanded *Pabpn1* allele are viable but mice with two expanded alleles (*Pabpn1*$^{A17/A17}$) are not. Further investigation is necessary to determine whether expanded PABPN1 dosage correlates with aggregate phenotypes. Higher levels of mutant protein in *Pabpn1*$^{A17/A17}$ mice could increase aggregation, depleting the soluble pool of PABPN1 protein. As expanded PABPN1 is partially functional, an increase in PABPN1 aggregation could

further reduce the functionality of expanded PABPN1 protein by sequestering available, partially functional protein in aggregates. Overall, exploiting these mouse models to investigate how expanded *Pabpn1* dosage modulates viability phenotypes could identify critical roles for PABPN1 in development, and perhaps uncover dysfunctions that occur in OPMD.

The *Pabpn1*$^{A17/\Delta}$ animals also provide a critical resource as they only express expanded *Pabpn1*. Because the alanine expansion in PABPN1 is subtle, with the most common mutation expanding 10 alanines to 13 alanines in the PABPN1 protein, it is technically challenging to molecularly distinguish wildtype from expanded *Pabpn1 in vivo.* Therefore, this tool allows us to study expanded *Pabpn1* without confounding wildtype *Pabpn1*, which is necessary as homozygous expanded *Pabpn1*$^{A17/A17}$ mice are not generated in Mendelian ratios. Therefore, molecular and biochemical studies can be conducted using tissues and primary muscle cultures from these *Pabpn1*$^{A17/\Delta}$ animals to specifically interrogate what phenotypes occur in animals that express only expanded PABPN1. Performing crosslinking-immunoprecipitation followed by sequencing (CLIP-seq) for expanded PABPN1 on muscle tissue or primary muscle cells from these animals is one example of a critical experiment that is now feasible. This experiment would help to identify what transcripts are targeted by the wildtype PABPN1 protein, and whether expanded PABPN1 binds to the same set of target RNAs. These data would also give us insight into whether wildtype PABPN1 is able to bind A-rich tracts within transcripts or if binding is exclusive to poly(A) tails, and if expanded PABPN1 has different binding preferences. Unpublished mass spectrometry analysis from our lab reveals that expanded PABPN1 protein co-immunoprecipitates more proteins than wildtype PABPN1 (195).

These CLIP-seq results would supplement these data, yielding interesting information as to whether the N-terminal expansion also modulates protein-transcript interactions, which is likely influenced by these protein-protein interactions. Many RNA binding proteins such as HuR interact more with expanded PABPN1 than with wildtype PABPN1 (195), which could influence where the PABPN1 protein is binding within transcripts as HuR is an AU-rich binding protein. It is also possible that these protein-protein interactions with expanded PABPN1 sequester proteins in aggregates, allowing other proteins that under wildtype conditions would not be able to compete with these factors to bind novel target RNAs. These data could also shed light on how PABPN1 regulates poly(A) site selection (PAS) as PABPN1 has been hypothesized to mask weak PASs (42), and whether expanded PABPN1 is able to fulfill this role in alternative polyadenylation. Overall, these mouse models provide critical tools that allow us to interrogate to what extent expanded PABPN1 is able to target and potentially regulate transcripts in muscle, and whether RNA phenotypes could contribute to muscle defects in OPMD.

In Chapter 3, we determined that HuR regulates *Pabpn1* transcript and protein levels. Importantly, we detect an ~1.5 fold increase in PABPN1 protein levels in $HuR^{\Delta/\Delta}$ muscles relative to control $HuR^{fl/fl}$ muscles. We hypothesize that this subtle increase in PABPN1 protein levels in $HuR^{\Delta/\Delta}$ muscles relative to control $HuR^{fl/fl}$ muscles could be protective against OPMD pathology. Overexpression of wildtype PABPN1 protein is protective against age-related dysphagia (65), suggesting that this approach is feasible and could modulate OPMD pathology. We could test this hypothesis by crossing heterozygous $Pabpn1^{A17/+}$ mice to $HuR^{\Delta/\Delta}$ mice to determine if HuR depletion rescues phenotypes observed with expanded PABPN1 such as bulk poly(A) tail length and

mitochondrial defects. Future work using unbiased high-throughput technologies to identify additional RNA binding proteins bound to the 3'UTR of the *Pabpn1* transcript is necessary as HuR is likely only a component of the post-transcriptional regulatory machinery involved in modulating *Pabpn1* expression in muscle. As PABPN1 protein levels are even lower in muscles affected in OPMD such as pharynx than other skeletal muscles (144), investigating whether HuR plays a more prominent role in *Pabpn1* regulation in these tissues, or if additional regulatory mechanisms are involved to further decrease PABPN1 levels is critical.

Several clinical trials are currently underway for OPMD treatments. One recently completed phase 2 clinical trial (NCT02015481) focused on the disaccharide trehalose, which reduces PABPN1 aggregation, muscle weakness, and the number of TUNEL-positive nuclei in an OPMD mouse model overexpressing expanded PABPN1(82). This clinical trial concluded that trehalose was safe and well-tolerated. However, this study was not powered for efficacy as only 25 patients were enrolled. Therefore, future work with larger patient cohorts is needed to determine the efficacy of trehalose treatment in OPMD.

Gene therapy approaches, which genetically modify disease-causing mutations, have also been proposed for OPMD treatment. Although gene therapies for OPMD are not currently in clinical trials, Benitec Biopharma is preparing to start human studies for an OPMD gene therapy in late 2018. This therapy, BB-301, is a vector containing DNA-directed RNA interference to knockdown expanded *PABPN1* while also encoding a wildtype *PABPN1* allele to replace the mutant gene. This approach has been shown to alleviate muscle phenotypes in a mouse model overexpressing expanded PABPN1

including insoluble aggregates and muscle weakness (114). Although this approach is promising, the knockdown/overexpression vector was introduced into specific mouse muscles via intramuscular injection (114). It is unclear how these future human studies intend to deliver vectors to OPMD-affected muscles. Intramuscular injections could be problematic considering that human muscles are much larger than mouse muscles and could result in heterogeneous knockdown and overexpression of *PABPN1* throughout affected muscles. Additionally, because OPMD often affects multiple muscles including pharynx and quadriceps, this delivery approach could require injecting multiple muscles, with potentially multiple injections required per muscle. However, several gene therapy approaches utilizing systemic delivery to target particular tissues, including skeletal muscles, have been successful. A subset of Duchenne Muscular Dystrophy (DMD) patients harbor mutations encoding premature stop codons in the critical muscle protein Dystrophin, and clinical trials testing exon-skipping drugs that result in truncated, partially functional Dystrophin protein are currently underway. For example, systemic delivery of eteplirson, an FDA-approved exon-skipping DMD drug developed by Sarepta Therapeutics, serves as a proof of principle that systemic drug delivery of drugs treating muscle diseases is feasible. Systemic delivery of therapies that target muscle-specific events, such muscle-specific proteins, avoid complications associated with muscle-specific delivery. However, because the mutations causing OPMD occur in a ubiquitous protein, PABPN1, systemic treatments are potentially problematic as they could cause effects in non-muscle tissues.

Ubiquitous expression of expanded PABPN1 causes muscle-specific phenotypes, demonstrating that non-muscle tissues are able to function normally when endogenous

levels of expanded PABPN1 are present. It is unknown, however, how systemic BB-301 treatment could affect non-muscle tissues as systemic knockdown of expanded *PABPN1* and overexpression of wildtype PABPN1 has not been studied in non-muscle tissues, and the requirements for PABPN1 levels in these tissues is unknown. Therefore, a gene therapy approach that takes advantage of mature skeletal muscle-specific characteristics could be ideal where systemic treatment could modulate *PABPN1* specifically in skeletal muscles. The work presented in Chapters 3 and 4 has identified potential targets for modulating *Pabpn1* expression in muscle, which could be protective against OPMD pathology. The therapeutic implications of work presented in Chapter 3 are critical as we have determined that HuR regulates PABPN1 levels specifically in myotubes, an *in vitro* model of skeletal muscle. Because we do not observe HuR-mediated *Pabpn1* regulation in myoblasts, this work suggests that modulating HuR-mediated *Pabpn1* regulation could be a method of specifically targeting OPMD therapies to mature skeletal muscle. One approach to develop novel therapies for OPMD would be designing antisense oligonucleotides that block HuR binding to the *Pabpn1* 3'UTR to increase PABPN1 protein levels in OPMD muscle. This finding is important because current investigations into OPMD therapies do not take advantage of approaches where systemic treatments could target muscle specifically. However, a complementary approach coupling BB-301 with an antisense oligonucleotide that blocks HuR binding to the *Pabpn1* 3'UTR could also be beneficial. For example, we could systemically introduce a vector containing an antisense oligonucleotide that selectively knocks down expanded *Pabpn1* and an antisense oligonucleotide that blocks HuR binding to the *Pabpn1* 3'UTR, with both antisense oligonucleotides placed under the control of a muscle-specific promoter. This

approach would avoid potential problems associated with modulating *Pabpn1* expression in non-muscle tissues, which do not require treatment, while also stimulating wildtype *Pabpn1* expression without exogenous expression.

Together, our work has identified several pathways that regulate PABPN1 levels in muscle, which could predispose this tissue to pathology when PABPN1 function is reduced in OPMD. These findings demonstrate that the complex interplay between post-transcriptional regulatory elements, such as RNA binding proteins and microRNAs, modulate *Pabpn1* transcript and protein levels in muscle, which could predispose this tissue to pathology when *PABPN1* is mutated in OPMD.

5.5 Final conclusions

Overall, the work presented in this dissertation has generated critical mouse models of OPMD and determined a potential pathway that could be modulated in OPMD therapies. Our work has focused on understanding how endogenous levels of expanded PABPN1 causes phenotypes at the molecular, cellular, and tissue-wide levels. As low PABPN1 levels could predispose muscle to OPMD pathology, we identified multiple post-transcriptional mechanisms that contribute to low PABPN1 levels specifically in muscle. Together, our work suggests that expansion mutations in the *PABPN1* gene cause a partial loss of function and that raising PABPN1 levels could be protective (Figure 5.2). This work is consistent with the model that loss of function could be due to the inability of expanded PABPN1 to participate in molecular events, perhaps through sequestration by nuclear aggregates and/or protein-inherent deficiencies. This work is significant

because we have generated critical mouse models and identified regulatory mechanisms that can be used to further interrogate the mechanisms contributing to OPMD pathology, perhaps uncovering therapeutic targets for OPMD treatment. Overall, our work serves as a paradigm for studying tissue-specific functions, and regulation, of RNA binding proteins. These studies are critical to understand why some tissues are susceptible to mutations in these genes to develop effective therapeutic approaches.

5.6 Figures

Legends to Figures

Figure 5.1: The complement of *trans*-acting factors bound to the *Pabpn1* 3'UTR could influence *Pabpn1* expression in a tissue-specific manner. Chapters 3 and 4 have identified several *trans*-acting factors, including the RNA binding protein HuR and microRNAs, that regulate *Pabpn1* expression in muscle cells (left panel). The complement of factors bound to the *Pabpn1* 3'UTR in non-muscle tissue (right panel) likely differs from the factors bound in muscle tissue, contributing to differences in PABPN1 protein levels across tissues.

Figure 5.2: Identification of loss of function phenotypes and mechanisms that contribute to low PABPN1 levels in muscle: Implications for OPMD treatment. We hypothesized that low PABPN1 levels could predispose muscle to OPMD pathology if mutated *PABPN1* is not fully functional. Expression of expanded PABPN1 could bring functional PABPN1 levels below a threshold for pathology only in muscle (top panel). The work presented in this dissertation has focused on investigating two main questions:

One, to what extent is expanded PABPN1 functional when expressed at endogenous levels (middle left panel); And two, what are the post-transcriptional mechanisms contributing to low PABPN1 protein levels specifically in muscle (middle right panel). Partial overlap of phenotypes observed *Pabpn1*$^{+/A17}$ and *Pabpn1*$^{+/\Delta}$ mice suggest that expanded PABPN1 is partially functional (middle left panel). We also determined that the RNA binding protein HuR and two microRNAs, miR-141-5p and miR-331-3p, negatively regulate *Pabpn1* expression in muscle. Together, this work suggests that modulating *Pabpn1* transcript and protein levels in muscle could raise functional PABPN1 protein levels above this threshold for pathology (bottom panel) and could be protective for OPMD patients.

Figure 5.1

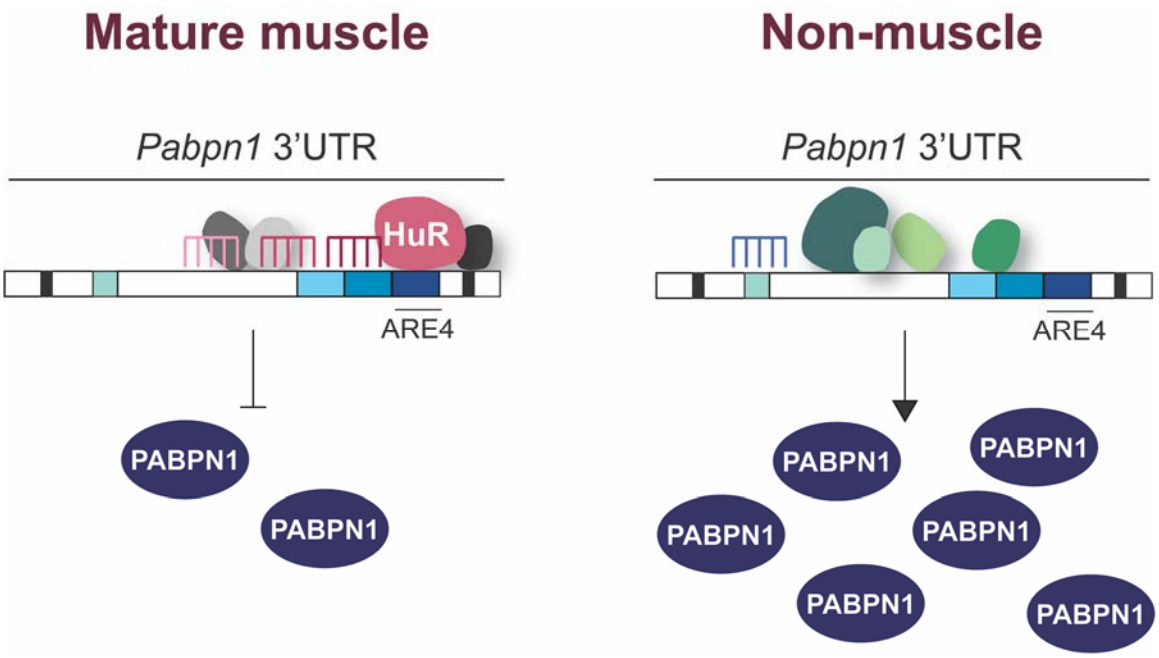
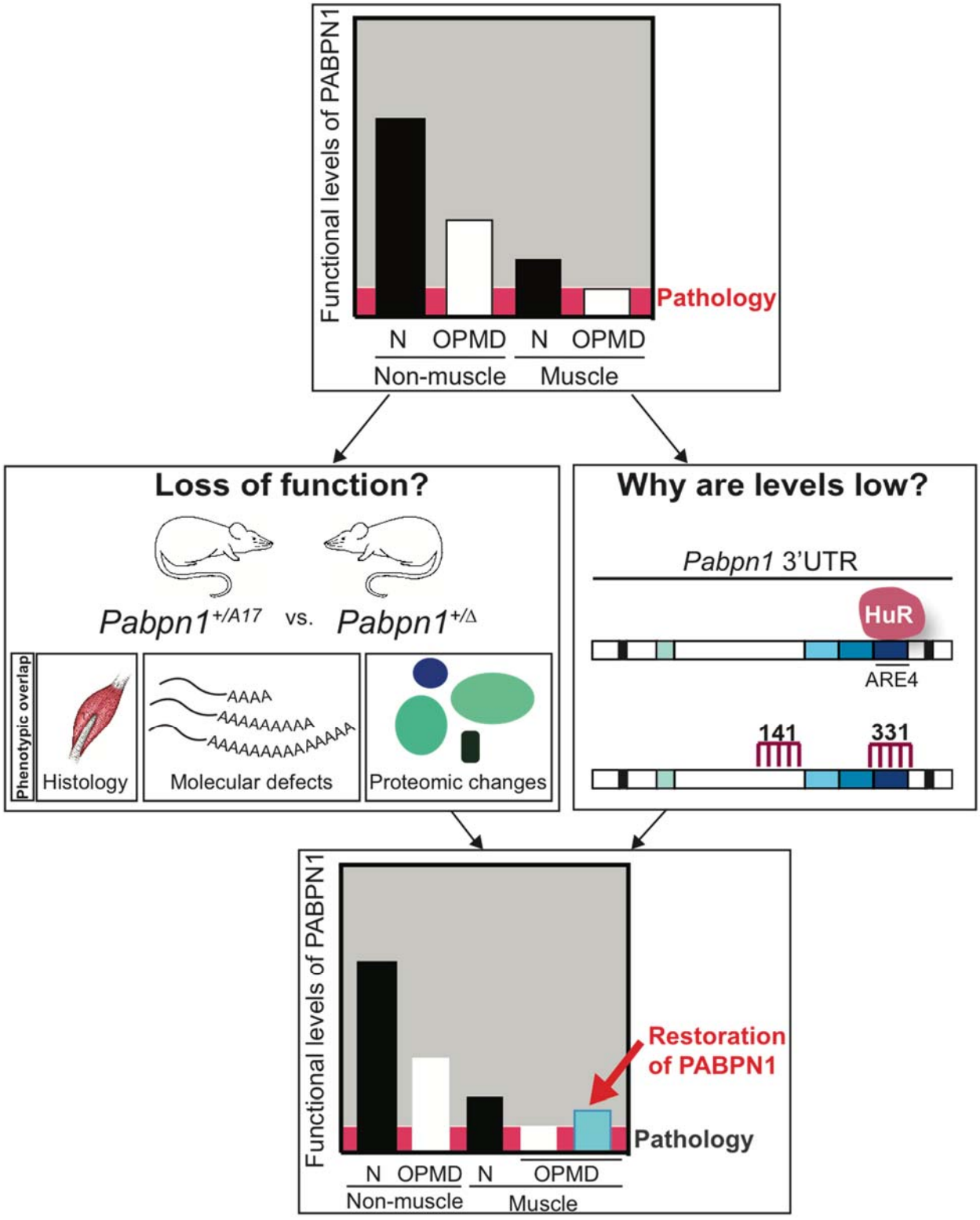

Mature muscle
Pabpn1 3'UTR
HuR
ARE4
PABPN1
PABPN1
Non-muscle
Pabpn1 3'UTR
ARE4
PABPN1
PABPN1
PABPN1
PABPN1
PABPN1
PABPN1

Figure 5.2



REFERENCES

1. Corbett, A.H. (2018) Post-transcriptional Regulation of Gene Expression and Human Disease. *Current Opinions in Cell Biology*, **52**.
2. Vest, K.E., Phillips, B.L., Banerjee, A., Apponi, L.H., Dammer, E.B., Xu, W., Zheng, D., Yu, J., Tian, B., Pavlath, G.K. *et al.* (2017) Novel mouse models of oculopharyngeal muscular dystrophy (OPMD) reveal early onset mitochondrial defects and suggest loss of PABPN1 may contribute to pathology. *Hum Mol Genet*, **26**, 3235-3252.
3. Ramanathan, A., Robb, G.B. and Chan, S.H. (2016) mRNA capping: biological functions and applications. *Nucleic Acids Res*, **44**, 7511-7526.
4. Shi, Y. (2017) Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol*, **18**, 655-670.
5. Hirose, Y. and Manley, J.L. (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev*, **14**, 1415-1429.
6. Proudfoot, N.J., Furger, A. and Dye, M.J. (2002) Integrating mRNA processing with transcription. *Cell*, **108**, 501-512.
7. McKee, A.E. and Silver, P.A. (2007) Systems perspectives on mRNA processing. *Cell Res*, **17**, 581-590.
8. Corden, J.L. (1990) Tails of RNA polymerase II. *Trends Biochem Sci*, **15**, 383-387.
9. Phatnani, H.P. and Greenleaf, A.L. (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*, **20**, 2922-2936.
10. Komarnitsky, P., Cho, E.J. and Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*, **14**, 2452-2460.
11. Morris, D.P., Michelotti, G.A. and Schwinn, D.A. (2005) Evidence that phosphorylation of the RNA polymerase II carboxyl-terminal repeats is similar in yeast and humans. *J Biol Chem*, **280**, 31368-31377.
12. Cho, E.J., Takagi, T., Moore, C.R. and Buratowski, S. (1997) mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev*, **11**, 3319-3326.
13. Cho, E.J., Rodriguez, C.R., Takagi, T. and Buratowski, S. (1998) Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes Dev*, **12**, 3482-3487.
14. Connelly, S. and Manley, J.L. (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev*, **2**, 440-452.
15. Whitelaw, E. and Proudfoot, N. (1986) Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. *EMBO J*, **5**, 2915-2922.
16. Mayr, C. (2017) Regulation by 3'-Untranslated Regions. *Annu Rev Genet*, **51**, 171-194.
17. Friedman, R.C., Farh, K.K., Burge, C.B. and Bartel, D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res*, **19**, 92-105.

18. Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S. and Kellis, M. (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, **434**, 338-345.
19. Xu, N., Chen, C.Y. and Shyu, A.B. (1997) Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay. *Mol Cell Biol*, **17**, 4611-4621.
20. Barreau, C., Paillard, L. and Osborne, H.B. (2005) AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res*, **33**, 7138-7150.
21. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P. and Burge, C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, **115**, 787-798.
22. Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-297.
23. Croce, C.M. and Calin, G.A. (2005) miRNAs, cancer, and stem cell division. *Cell*, **122**, 6-7.
24. Naguibneva, I., Ameyar-Zazoua, M., Polesskaya, A., Ait-Si-Ali, S., Groisman, R., Souidi, M., Cuvellier, S. and Harel-Bellan, A. (2006) The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. *Nat Cell Biol*, **8**, 278-284.
25. He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D. *et al.* (2007) A microRNA component of the p53 tumour suppressor network. *Nature*, **447**, 1130-1134.
26. Cogswell, J.P., Ward, J., Taylor, I.A., Waters, M., Shi, Y., Cannon, B., Kelnar, K., Kemppainen, J., Brown, D., Chen, C. *et al.* (2008) Identification of miRNA changes in Alzheimer's disease brain and CSF yields putative biomarkers and insights into disease pathways. *J Alzheimers Dis*, **14**, 27-41.
27. Johnson, R., Zuccato, C., Belyaev, N.D., Guest, D.J., Cattaneo, E. and Buckley, N.J. (2008) A microRNA-based gene dysregulation pathway in Huntington's disease. *Neurobiol Dis*, **29**, 438-445.
28. Vinciguerra, P. and Stutz, F. (2004) mRNA export: an assembly line from genes to nuclear pores. *Curr Opin Cell Biol*, **16**, 285-292.
29. Garber, K., Smith, K.T., Reines, D. and Warren, S.T. (2006) Transcription, translation and fragile X syndrome. *Curr Opin Genet Dev*, **16**, 270-275.
30. Chen, E., Sharma, M.R., Shi, X., Agrawal, R.K. and Joseph, S. (2014) Fragile X mental retardation protein regulates translation by binding directly to the ribosome. *Mol Cell*, **54**, 407-417.
31. Cammas, A., Sanchez, B.J., Lian, X.J., Dormoy-Raclet, V., van der Giessen, K., Lopez de Silanes, I., Ma, J., Wilusz, C., Richardson, J., Gorospe, M. *et al.* (2014) Destabilization of nucleophosmin mRNA by the HuR/KSRP complex is required for muscle fibre formation. *Nat Commun*, **5**, 4190.
32. Singh, G., Pratt, G., Yeo, G.W. and Moore, M.J. (2015) The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annu Rev Biochem*, **84**, 325-354.
33. Banerjee, A., Apponi, L.H., Pavlath, G.K. and Corbett, A.H. (2013) PABPN1: molecular function and muscle disease. *FEBS J*, **280**, 4230-4250.

34. Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A. *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, **499**, 172-177.
35. Taliaferro, J.M., Lambert, N.J., Sudmant, P.H., Dominguez, D., Merkin, J.J., Alexis, M.S., Bazile, C. and Burge, C.B. (2016) RNA Sequence Context Effects Measured In Vitro Predict In Vivo Protein Binding and Regulation. *Mol Cell*, **64**, 294-306.
36. Gerstberger, S., Hafner, M. and Tuschl, T. (2014) A census of human RNA-binding proteins. *Nat Rev Genet*, **15**, 829-845.
37. Harish, P., Malerba, A., Dickson, G. and Bachtarzi, H. (2015) Progress on gene therapy, cell therapy, and pharmacological strategies toward the treatment of oculopharyngeal muscular dystrophy. *Hum Gene Ther*, **26**, 286-292.
38. Brais, B., Bouchard, J.P., Xie, Y.G., Rochefort, D.L., Chretien, N., Tome, F.M., Lafreniere, R.G., Rommens, J.M., Uyama, E., Nohira, O. *et al.* (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet*, **18**, 164-167.
39. Beaulieu, Y.B., Kleinman, C.L., Landry-Voyer, A.M., Majewski, J. and Bachand, F. (2012) Polyadenylation-dependent control of long noncoding RNA expression by the poly(A)-binding protein nuclear 1. *PLoS Genet*, **8**, e1003078.
40. Bresson, S.M., Hunter, O.V., Hunter, A.C. and Conrad, N.K. (2015) Canonical Poly(A) Polymerase Activity Promotes the Decay of a Wide Variety of Mammalian Nuclear RNAs. *PLoS Genet*, **11**, e1005610.
41. de Klerk, E., Venema, A., Anvar, S.Y., Goeman, J.J., Hu, O., Trollet, C., Dickson, G., den Dunnen, J.T., van der Maarel, S.M., Raz, V. *et al.* (2012) Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. *Nucleic Acids Res*, **40**, 9089-9101.
42. Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kuhn, U., Menzies, F.M., Oude Vrielink, J.A., Bos, A.J., Drost, J., Rooijers, K. *et al.* (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell*, **149**, 538-553.
43. Lemay, J.F., D'Amours, A., Lemieux, C., Lackner, D.H., St-Sauveur, V.G., Bahler, J. and Bachand, F. (2010) The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. *Mol Cell*, **37**, 34-45.
44. Li, W., You, B., Hoque, M., Zheng, D., Luo, W., Ji, Z., Park, J.Y., Gunderson, S.I., Kalsotra, A., Manley, J.L. *et al.* (2015) Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet*, **11**, e1005166.
45. Kuhn, U., Nemeth, A., Meyer, S. and Wahle, E. (2003) The RNA binding domains of the nuclear poly(A)-binding protein. *J Biol Chem*, **278**, 16916-16925.
46. Wahle, E. (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell*, **66**, 759-768.
47. Wahle, E., Lustig, A., Jeno, P. and Maurer, P. (1993) Mammalian poly(A)-binding protein II. Physical properties and binding to polynucleotides. *J Biol Chem*, **268**, 2937-2945.

48. Keller, W., Bienroth, S., Lang, K.M. and Christofori, G. (1991) Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J*, **10**, 4241-4249.
49. MacDonald, C.C., Wilusz, J. and Shenk, T. (1994) The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol Cell Biol*, **14**, 6647-6654.
50. Apponi, L.H., Leung, S.W., Williams, K.R., Valentini, S.R., Corbett, A.H. and Pavlath, G.K. (2010) Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet*, **19**, 1058-1065.
51. Elkon, R., Ugalde, A.P. and Agami, R. (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet*, **14**, 496-506.
52. Muniz, L., Davidson, L. and West, S. (2015) Poly(A) Polymerase and the Nuclear Poly(A) Binding Protein, PABPN1, Coordinate the Splicing and Degradation of a Subset of Human Pre-mRNAs. *Mol Cell Biol*, **35**, 2218-2230.
53. Ishigaki, Y., Li, X., Serin, G. and Maquat, L.E. (2001) Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell*, **106**, 607-617.
54. Lemieux, C. and Bachand, F. (2009) Cotranscriptional recruitment of the nuclear poly(A)-binding protein Pab2 to nascent transcripts and association with translating mRNPs. *Nucleic Acids Res*, **37**, 3418-3430.
55. Bresson, S.M. and Conrad, N.K. (2013) The human nuclear poly(a)-binding protein promotes RNA hyperadenylation and decay. *PLoS Genet*, **9**, e1003893.
56. Davies, J.E., Wang, L., Garcia-Oroz, L., Cook, L.J., Vacher, C., O'Donovan, D.G. and Rubinsztein, D.C. (2005) Doxycycline attenuates and delays toxicity of the oculopharyngeal muscular dystrophy mutation in transgenic mice. *Nat Med*, **11**, 672-677.
57. Winter, R., Kuhn, U., Hause, G. and Schwarz, E. (2012) Polyalanine-independent conformational conversion of nuclear poly(A)-binding protein 1 (PABPN1). *J Biol Chem*, **287**, 22662-22671.
58. Apponi, L.H., Corbett, A.H. and Pavlath, G.K. (2013) Control of mRNA stability contributes to low levels of nuclear poly(A) binding protein 1 (PABPN1) in skeletal muscle. *Skelet Muscle*, **3**, 23.
59. Anvar, S.Y., Raz, Y., Verway, N., van der Sluijs, B., Venema, A., Goeman, J.J., Vissing, J., van der Maarel, S.M., t Hoen, P.A., van Engelen, B.G. *et al.* (2013) A decline in PABPN1 induces progressive muscle weakness in oculopharyngeal muscle dystrophy and in muscle aging. *Aging (Albany NY)*, **5**, 412-426.
60. Kusters, B., van Hoeve, B.J., Schelhaas, H.J., Ter Laak, H., van Engelen, B.G. and Lammens, M. (2009) TDP-43 accumulation is common in myopathies with rimmed vacuoles. *Acta Neuropathol*, **117**, 209-211.
61. Fan, X., Messaed, C., Dion, P., Laganiere, J., Brais, B., Karpati, G. and Rouleau, G.A. (2003) HnRNP A1 and A/B interaction with PABPN1 in oculopharyngeal muscular dystrophy. *Can J Neurol Sci*, **30**, 244-251.
62. Klein, P., Oloko, M., Roth, F., Montel, V., Malerba, A., Jarmin, S., Gidaro, T., Popplewell, L., Perie, S., Lacau St Guily, J. *et al.* (2016) Nuclear poly(A)-binding protein aggregates misplace a pre-mRNA outside of SC35 speckle causing its abnormal splicing. *Nucleic Acids Res*, **44**, 10929-10945.

63. Riaz, M., Raz, Y., van Putten, M., Paniagua-Soriano, G., Krom, Y.D., Florea, B.I. and Raz, V. (2016) PABPN1-Dependent mRNA Processing Induces Muscle Wasting. *PLoS Genet*, **12**, e1006031.
64. Davies, J.E., Sarkar, S. and Rubinsztein, D.C. (2008) Wild-type PABPN1 is anti-apoptotic and reduces toxicity of the oculopharyngeal muscular dystrophy mutation. *Hum Mol Genet*, **17**, 1097-1108.
65. Randolph, M.E., Luo, Q., Ho, J., Vest, K.E., Sokoloff, A.J. and Pavlath, G.K. (2014) Ageing and muscular dystrophy differentially affect murine pharyngeal muscles in a region-dependent manner. *J Physiol*, **592**, 5301-5315.
66. Victor, M., Hayes, R. and Adams, R.D. (1962) Oculopharyngeal muscular dystrophy. A familial disease of late life characterized by dysphagia and progressive ptosis of the evelids. *N Engl J Med*, **267**, 1267-1272.
67. Youssof, S. (2016) The relationship between physical symptoms and health-related quality of life in oculopharyngeal muscular dystrophy. *Muscle Nerve*, **53**, 694-699.
68. Richard, P., Trollet, C., Gidaro, T., Demay, L., Brochier, G., Malfatti, E., Tom, F.M., Fardeau, M., Lafor, P., Romero, N. *et al.* (2015) PABPN1 (GCN)11 as a Dominant Allele in Oculopharyngeal Muscular Dystrophy -Consequences in Clinical Diagnosis and Genetic Counselling. *J Neuromuscul Dis*, **2**, 175-180.
69. Jouan, L., Rocheford, D., Szuto, A., Carney, E., David, K., Dion, P.A. and Rouleau, G.A. (2014) An 18 alanine repeat in a severe form of oculopharyngeal muscular dystrophy. *Can J Neurol Sci*, **41**, 508-511.
70. Kerwitz, Y., Kuhn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E. and Wahle, E. (2003) Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J*, **22**, 3705-3714.
71. Kuhn, U., Gundel, M., Knoth, A., Kerwitz, Y., Rudel, S. and Wahle, E. (2009) Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J Biol Chem*, **284**, 22803-22814.
72. Wahle, E. (1995) Poly(A) tail length control is caused by termination of processive synthesis. *J Biol Chem*, **270**, 2800-2808.
73. Tome, F.M. and Fardeau, M. (1980) Nuclear inclusions in oculopharyngeal dystrophy. *Acta Neuropathol*, **49**, 85-87.
74. Corbeil-Girard, L.P., Klein, A.F., Sasseville, A.M., Lavoie, H., Dicaire, M.J., Saint-Denis, A., Page, M., Duranceau, A., Codere, F., Bouchard, J.P. *et al.* (2005) PABPN1 overexpression leads to upregulation of genes encoding nuclear proteins that are sequestered in oculopharyngeal muscular dystrophy nuclear inclusions. *Neurobiol Dis*, **18**, 551-567.
75. Tavanez, J.P., Bengoechea, R., Berciano, M.T., Lafarga, M., Carmo-Fonseca, M. and Enguita, F.J. (2009) Hsp70 chaperones and type I PRMTs are sequestered at intranuclear inclusions caused by polyalanine expansions in PABPN1. *PLoS One*, **4**, e6418.
76. Gidaro, T., Negroni, E., Perie, S., Mirabella, M., Laine, J., Lacau St Guily, J., Butler-Browne, G., Mouly, V. and Trollet, C. (2013) Atrophy, fibrosis, and

increased PAX7-positive cells in pharyngeal muscles of oculopharyngeal muscular dystrophy patients. *J Neuropathol Exp Neurol*, **72**, 234-243.
77. Uyama, E., Tsukahara, T., Goto, K., Kurano, Y., Ogawa, M., Kim, Y.J., Uchino, M. and Arahata, K. (2000) Nuclear accumulation of expanded PABP2 gene product in oculopharyngeal muscular dystrophy. *Muscle Nerve*, **23**, 1549-1554.
78. Dion, P., Shanmugam, V., Gaspar, C., Messaed, C., Meijer, I., Toulouse, A., Laganiere, J., Roussel, J., Rochefort, D., Laganiere, S. *et al.* (2005) Transgenic expression of an expanded (GCG)13 repeat PABPN1 leads to weakness and coordination defects in mice. *Neurobiol Dis*, **18**, 528-536.
79. Berciano, M.T., Villagra, N.T., Ojeda, J.L., Navascues, J., Gomes, A., Lafarga, M. and Carmo-Fonseca, M. (2004) Oculopharyngeal muscular dystrophy-like nuclear inclusions are present in normal magnocellular neurosecretory neurons of the hypothalamus. *Hum Mol Genet*, **13**, 829-838.
80. Villagra, N.T., Bengoechea, R., Vaque, J.P., Llorca, J., Berciano, M.T. and Lafarga, M. (2008) Nuclear compartmentalization and dynamics of the poly(A)-binding protein nuclear 1 (PABPN1) inclusions in supraoptic neurons under physiological and osmotic stress conditions. *Mol Cell Neurosci*, **37**, 622-633.
81. Davies, J.E., Rose, C., Sarkar, S. and Rubinsztein, D.C. (2010) Cystamine suppresses polyalanine toxicity in a mouse model of oculopharyngeal muscular dystrophy. *Sci Transl Med*, **2**, 34ra40.
82. Davies, J.E., Sarkar, S. and Rubinsztein, D.C. (2006) Trehalose reduces aggregate formation and delays pathology in a transgenic mouse model of oculopharyngeal muscular dystrophy. *Hum Mol Genet*, **15**, 23-31.
83. Brais, B., Xie, Y.G., Sanson, M., Morgan, K., Weissenbach, J., Korczyn, A.D., Blumen, S.C., Fardeau, M., Tome, F.M., Bouchard, J.P. *et al.* (1995) The oculopharyngeal muscular dystrophy locus maps to the region of the cardiac alpha and beta myosin heavy chain genes on chromosome 14q11.2-q13. *Hum Mol Genet*, **4**, 429-434.
84. Wang, Q. and Bag, J. (2006) Ectopic expression of a polyalanine expansion mutant of poly(A)-binding protein N1 in muscle cells in culture inhibits myogenesis. *Biochem Biophys Res Commun*, **340**, 815-822.
85. Raz, V., Routledge, S., Venema, A., Buijze, H., van der Wal, E., Anvar, S., Straasheijm, K.R., Klooster, R., Antoniou, M. and van der Maarel, S.M. (2011) Modeling oculopharyngeal muscular dystrophy in myotube cultures reveals reduced accumulation of soluble mutant PABPN1 protein. *Am J Pathol*, **179**, 1988-2000.
86. Catoire, H., Pasco, M.Y., Abu-Baker, A., Holbert, S., Tourette, C., Brais, B., Rouleau, G.A., Parker, J.A. and Neri, C. (2008) Sirtuin inhibition protects from the polyalanine muscular dystrophy protein PABPN1. *Hum Mol Genet*, **17**, 2108-2117.
87. Chartier, A., Benoit, B. and Simonelig, M. (2006) A Drosophila model of oculopharyngeal muscular dystrophy reveals intrinsic toxicity of PABPN1. *EMBO J*, **25**, 2253-2262.
88. Hino, H., Araki, K., Uyama, E., Takeya, M., Araki, M., Yoshinobu, K., Miike, K., Kawazoe, Y., Maeda, Y., Uchino, M. *et al.* (2004) Myopathy phenotype in

transgenic mice expressing mutated PABPN1 as a model of oculopharyngeal muscular dystrophy. *Hum Mol Genet*, **13**, 181-190.
89. Mankodi, A., Wheeler, T.M., Shetty, R., Salceies, K.M., Becher, M.W. and Thornton, C.A. (2012) Progressive myopathy in an inducible mouse model of oculopharyngeal muscular dystrophy. *Neurobiol Dis*, **45**, 539-546.
90. Chartier, A., Klein, P., Pierson, S., Barbezier, N., Gidaro, T., Casas, F., Carberry, S., Dowling, P., Maynadier, L., Bellec, M. *et al.* (2015) Mitochondrial dysfunction reveals the role of mRNA poly(A) tail regulation in oculopharyngeal muscular dystrophy pathogenesis. *PLoS Genet*, **11**, e1005092.
91. Anvar, S.Y., t Hoen, P.A., Venema, A., van der Sluijs, B., van Engelen, B., Snoeck, M., Vissing, J., Trollet, C., Dickson, G., Chartier, A. *et al.* (2011) Deregulation of the ubiquitin-proteasome system is the predominant molecular pathology in OPMD animal models and patients. *Skelet Muscle*, **1**, 15.
92. Trollet, C., Anvar, S.Y., Venema, A., Hargreaves, I.P., Foster, K., Vignaud, A., Ferry, A., Negroni, E., Hourde, C., Baraibar, M.A. *et al.* (2010) Molecular and phenotypic characterization of a mouse model of oculopharyngeal muscular dystrophy reveals severe muscular atrophy restricted to fast glycolytic fibres. *Hum Mol Genet*, **19**, 2191-2207.
93. Scheuermann, T., Schulz, B., Blume, A., Wahle, E., Rudolph, R. and Schwarz, E. (2003) Trinucleotide expansions leading to an extended poly-L-alanine segment in the poly (A) binding protein PABPN1 cause fibril formation. *Protein Sci*, **12**, 2685-2692.
94. Lakso, M., Pichel, J.G., Gorman, J.R., Sauer, B., Okamoto, Y., Lee, E., Alt, F.W. and Westphal, H. (1996) Efficient in vivo manipulation of mouse genomic sequences at the zygote stage. *Proc Natl Acad Sci U S A*, **93**, 5860-5865.
95. Vest, K.E., Apponi, L.H., Banerjee, A., Pavlath, G.K. and Corbett, A.H. (2015) An Antibody to Detect Alanine-Expanded PABPN1: A New Tool to Study Oculopharyngeal Muscular Dystrophy. *J Neuromuscul Dis*, **2**, 439-446.
96. Cooper, R.N., Tajbakhsh, S., Mouly, V., Cossu, G., Buckingham, M. and Butler-Browne, G.S. (1999) In vivo satellite cell activation via Myf5 and MyoD in regenerating mouse skeletal muscle. *J Cell Sci*, **112 (Pt 17)**, 2895-2901.
97. Chekanova, J.A., Shaw, R.J. and Belostotsky, D.A. (2001) Analysis of an essential requirement for the poly(A) binding protein function using cross-species complementation. *Curr Biol*, **11**, 1207-1214.
98. Hoque, M., Ji, Z., Zheng, D., Luo, W., Li, W., You, B., Park, J.Y., Yehia, G. and Tian, B. (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods*, **10**, 133-139.
99. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, **4**, 44-57.
100. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, **37**, 1-13.
101. Cho, Y., Hazen, B.C., Gandra, P.G., Ward, S.R., Schenk, S., Russell, A.P. and Kralli, A. (2016) Perm1 enhances mitochondrial biogenesis, oxidative capacity, and fatigue resistance in adult skeletal muscle. *FASEB J*, **30**, 674-687.

102. Barrientos, A., Casademont, J., Cardellach, F., Estivill, X., Urbano-Marquez, A. and Nunes, V. (1997) Reduced steady-state levels of mitochondrial RNA and increased mitochondrial DNA amount in human brain with aging. *Brain Res Mol Brain Res*, **52**, 284-289.
103. Richard, P., Trollet, C., Stojkovic, T., de Becdelievre, A., Perie, S., Pouget, J., Eymard, B., Neurologists of French Neuromuscular Reference Centers, C. and Filnemus. (2017) Correlation between PABPN1 genotype and disease severity in oculopharyngeal muscular dystrophy. *Neurology*, **88**, 359-365.
104. Van Der Sluijs, B.M., Hoefsloot, L.H., Padberg, G.W., Van Der Maarel, S.M. and Van Engelen, B.G. (2003) Oculopharyngeal muscular dystrophy with limb girdle weakness as major complaint. *J Neurol*, **250**, 1307-1312.
105. Youssof, S., Schrader, R., Bear, D. and Morrison, L. (2015) Hip flexion weakness is associated with impaired mobility in oculopharyngeal muscular dystrophy: a retrospective study with implications for trial design. *Neuromuscul Disord*, **25**, 238-246.
106. Duddy, W., Duguez, S., Johnston, H., Cohen, T.V., Phadke, A., Gordish-Dressman, H., Nagaraju, K., Gnocchi, V., Low, S. and Partridge, T. (2015) Muscular dystrophy in the mdx mouse is a severe myopathy compounded by hypotrophy, hypertrophy and hyperplasia. *Skelet Muscle*, **5**, 16.
107. Tanabe, Y., Esaki, K. and Nomura, T. (1986) Skeletal muscle pathology in X chromosome-linked muscular dystrophy (mdx) mouse. *Acta Neuropathol*, **69**, 91-95.
108. Kelleher, A.R., Pereira, S.L., Jefferson, L.S. and Kimball, S.R. (2015) REDD2 expression in rat skeletal muscle correlates with nutrient-induced activation of mTORC1: responses to aging, immobilization, and remobilization. *Am J Physiol Endocrinol Metab*, **308**, E122-129.
109. Miyazaki, M. and Esser, K.A. (2009) REDD2 is enriched in skeletal muscle and inhibits mTOR signaling in response to leucine and stretch. *Am J Physiol Cell Physiol*, **296**, C583-592.
110. Gambelli, S., Malandrini, A., Ginanneschi, F., Berti, G., Cardaioli, E., De Stefano, R., Franci, M., Salvadori, C., Mari, F., Bruttini, M. *et al.* (2004) Mitochondrial abnormalities in genetically assessed oculopharyngeal muscular dystrophy. *Eur Neurol*, **51**, 144-147.
111. Sadeh, M., Pauzner, R., Blatt, I., Mouallem, M. and Farfel, Z. (1993) Mitochondrial abnormalities in oculopharyngeal muscular dystrophy. *Muscle Nerve*, **16**, 982-983.
112. Schroder, J.M., Krabbe, B. and Weis, J. (1995) Oculopharyngeal muscular dystrophy: clinical and morphological follow-up study reveals mitochondrial alterations and unique nuclear inclusions in a severe autosomal recessive type. *Neuropathol Appl Neurobiol*, **21**, 68-73.
113. Wong, K.T., Dick, D. and Anderson, J.R. (1996) Mitochondrial abnormalities in oculopharyngeal muscular dystrophy. *Neuromuscul Disord*, **6**, 163-166.
114. Malerba, A., Klein, P., Bachtarzi, H., Jarmin, S.A., Cordova, G., Ferry, A., Strings, V., Espinoza, M.P., Mamchaoui, K., Blumen, S.C. *et al.* (2017) PABPN1 gene therapy for oculopharyngeal muscular dystrophy. *Nat Commun*, **8**, 14848.

115. Kepp, K.P. (2016) Alzheimer's disease due to loss of function: A new synthesis of the available data. *Prog Neurobiol*, **143**, 36-60.
116. Arteaga-Bracho, E.E., Gulinello, M., Winchester, M.L., Pichamoorthy, N., Petronglo, J.R., Zambrano, A.D., Inocencio, J., De Jesus, C.D., Louie, J.O., Gokhan, S. *et al.* (2016) Postnatal and adult consequences of loss of huntingtin during development: Implications for Huntington's disease. *Neurobiol Dis*, **96**, 144-155.
117. Messaed, C., Dion, P.A., Abu-Baker, A., Rochefort, D., Laganiere, J., Brais, B. and Rouleau, G.A. (2007) Soluble expanded PABPN1 promotes cell death in oculopharyngeal muscular dystrophy. *Neurobiol Dis*, **26**, 546-557.
118. Kontgen, F., Suss, G., Stewart, C., Steinmetz, M. and Bluethmann, H. (1993) Targeted disruption of the MHC class II Aa gene in C57BL/6 mice. *Int Immunol*, **5**, 957-964.
119. van der Weyden, L., Adams, D.J., Harris, L.W., Tannahill, D., Arends, M.J. and Bradley, A. (2005) Null and conditional semaphorin 3B alleles using a flexible puroDeltatk loxP/FRT vector. *Genesis*, **41**, 171-178.
120. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408.
121. Ross, J.M. (2011) Visualization of mitochondrial respiratory function using cytochrome c oxidase/succinate dehydrogenase (COX/SDH) double-labeling histochemistry. *J Vis Exp*, e3266.
122. Spinazzi, M., Casarin, A., Pertegato, V., Salviati, L. and Angelini, C. (2012) Assessment of mitochondrial respiratory chain enzymatic activities on tissues and cultured cells. *Nat Protoc*, **7**, 1235-1246.
123. Zheng, D., Liu, X. and Tian, B. (2016) 3'READS+, a sensitive and accurate method for 3' end sequencing of polyadenylated RNA. *RNA*, **22**, 1631-1639.
124. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, **9**, 357-359.
125. Dammer, E.B., Lee, A.K., Duong, D.M., Gearing, M., Lah, J.J., Levey, A.I. and Seyfried, N.T. (2015) Quantitative phosphoproteomics of Alzheimer's disease reveals cross-talk between kinases and small heat shock proteins. *Proteomics*, **15**, 508-519.
126. Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J. and Klenk, D.C. (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem*, **150**, 76-85.
127. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N. and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*, **13**, 2513-2526.
128. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V. and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*, **10**, 1794-1805.
129. Kechavarzi, B. and Janga, S.C. (2014) Dissecting the expression landscape of RNA-binding proteins in human cancers. *Genome Biol*, **15**, R14.

130. Liu, X., Hoque, M., Larochelle, M., Lemay, J.F., Yurko, N., Manley, J.L., Bachand, F. and Tian, B. (2017) Comparative analysis of alternative polyadenylation in S. cerevisiae and S. pombe. *Genome Res*, **27**, 1685-1695.
131. Blau, H.M., Pavlath, G.K., Hardeman, E.C., Chiu, C.P., Silberstein, L., Webster, S.G., Miller, S.C. and Webster, C. (1985) Plasticity of the differentiated state. *Science*, **230**, 758-766.
132. Lyons, G.E., Ontell, M., Cox, R., Sassoon, D. and Buckingham, M. (1990) The expression of myosin genes in developing skeletal muscle in the mouse embryo. *J Cell Biol*, **111**, 1465-1476.
133. Jainchill, J.L., Aaronson, S.A. and Todaro, G.J. (1969) Murine sarcoma and leukemia viruses: assay using clonal lines of contact-inhibited mouse cells. *J Virol*, **4**, 549-553.
134. Reich, E., Franklin, R.M., Shatkin, A.J. and Tatum, E.L. (1961) Effect of actinomycin D on cellular nucleic acid synthesis and virus production. *Science*, **134**, 556-557.
135. Lee, J.E., Lee, J.Y., Wilusz, J., Tian, B. and Wilusz, C.J. (2010) Systematic analysis of cis-elements in unstable mRNAs demonstrates that CUGBP1 is a key regulator of mRNA decay in muscle cells. *PLoS One*, **5**, e11201.
136. Herrick, D.J. and Ross, J. (1994) The half-life of c-myc mRNA in growing and serum-stimulated cells: influence of the coding and 3' untranslated regions and role of ribosome translocation. *Mol Cell Biol*, **14**, 2119-2128.
137. Proudfoot, N.J. and Brownlee, G.G. (1976) 3' non-coding region sequences in eukaryotic messenger RNA. *Nature*, **263**, 211-214.
138. Fitzgerald, M. and Shenk, T. (1981) The sequence 5'-AAUAAA-3'forms parts of the recognition site for polyadenylation of late SV40 mRNAs. *Cell*, **24**, 251-260.
139. Wickens, M. and Stephenson, P. (1984) Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. *Science*, **226**, 1045-1051.
140. Chen, C.Y. and Shyu, A.B. (1994) Selective degradation of early-response-gene mRNAs: functional analyses of sequence features of the AU-rich elements. *Mol Cell Biol*, **14**, 8471-8482.
141. Peng, S.S., Chen, C.Y. and Shyu, A.B. (1996) Functional characterization of a non-AUUUA AU-rich element from the c-jun proto-oncogene mRNA: evidence for a novel class of AU-rich elements. *Mol Cell Biol*, **16**, 1490-1499.
142. Nagaoka, K., Suzuki, T., Kawano, T., Imakawa, K. and Sakai, S. (2006) Stability of casein mRNA is ensured by structural interactions between the 3'-untranslated region and poly(A) tail via the HuR and poly(A)-binding protein complex. *Biochim Biophys Acta*, **1759**, 132-140.
143. Chen, C.Y. and Shyu, A.B. (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem Sci*, **20**, 465-470.
144. Apponi, L.H., Corbett, A.H. and Pavlath, G.K. (2011) RNA-binding proteins and gene regulation in myogenesis. *Trends Pharmacol Sci*, **32**, 652-658.
145. Ma, W.J., Chung, S. and Furneaux, H. (1997) The Elav-like proteins bind to AU-rich elements and to the poly(A) tail of mRNA. *Nucleic Acids Res*, **25**, 3564-3569.

146. Lebedeva, S., Jens, M., Theil, K., Schwanhausser, B., Selbach, M., Landthaler, M. and Rajewsky, N. (2011) Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol Cell*, **43**, 340-352.
147. Mukherjee, N., Corcoran, D.L., Nusbaum, J.D., Reid, D.W., Georgiev, S., Hafner, M., Ascano, M., Jr., Tuschl, T., Ohler, U. and Keene, J.D. (2011) Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol Cell*, **43**, 327-339.
148. Fan, X.C. and Steitz, J.A. (1998) Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs. *EMBO J*, **17**, 3448-3460.
149. Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M. and Chen, C.Y. (2004) A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol Cell*, **14**, 571-583.
150. Zhang, L., Lee, J.E., Wilusz, J. and Wilusz, C.J. (2008) The RNA-binding protein CUGBP1 regulates stability of tumor necrosis factor mRNA in muscle cells: implications for myotonic dystrophy. *J Biol Chem*, **283**, 22457-22463.
151. Edwards, J.M., Long, J., de Moor, C.H., Emsley, J. and Searle, M.S. (2013) Structural insights into the targeting of mRNA GU-rich elements by the three RRMs of CELF1. *Nucleic Acids Res*, **41**, 7153-7166.
152. Masuda, A., Andersen, H.S., Doktor, T.K., Okamoto, T., Ito, M., Andresen, B.S. and Ohno, K. (2012) CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. *Sci Rep*, **2**, 209.
153. Katsanou, V., Milatos, S., Yiakouvaki, A., Sgantzis, N., Kotsoni, A., Alexiou, M., Harokopos, V., Aidinis, V., Hemberger, M. and Kontoyiannis, D.L. (2009) The RNA-binding protein Elavl1/HuR is essential for placental branching morphogenesis and embryonic development. *Mol Cell Biol*, **29**, 2762-2776.
154. Sanchez, B.J., Tremblay, A., Hall, D., Kovacs, E., Ma, J. F., Hallauer, P. L., Di Marco, S., Pause, A., Hastings, K., Gallouzi, I. E. (2018), *Submitted*.
155. Fan, X.C. and Steitz, J.A. (1998) HNS, a nuclear-cytoplasmic shuttling sequence in HuR. *Proc Natl Acad Sci U S A*, **95**, 15293-15298.
156. Figueroa, A., Cuadrado, A., Fan, J., Atasoy, U., Muscat, G.E., Munoz-Canoves, P., Gorospe, M. and Munoz, A. (2003) Role of HuR in skeletal myogenesis through coordinate regulation of muscle differentiation genes. *Mol Cell Biol*, **23**, 4991-5004.
157. von Roretz, C., Beauchamp, P., Di Marco, S. and Gallouzi, I.E. (2011) HuR and myogenesis: being in the right place at the right time. *Biochim Biophys Acta*, **1813**, 1663-1667.
158. Beauchamp, P., Nassif, C., Hillock, S., van der Giessen, K., von Roretz, C., Jasmin, B.J. and Gallouzi, I.E. (2010) The cleavage of HuR interferes with its transportin-2-mediated nuclear import and promotes muscle fiber formation. *Cell Death Differ*, **17**, 1588-1599.
159. van der Giessen, K., Di-Marco, S., Clair, E. and Gallouzi, I.E. (2003) RNAi-mediated HuR depletion leads to the inhibition of muscle cell differentiation. *J Biol Chem*, **278**, 47119-47128.

160. van der Giessen, K. and Gallouzi, I.E. (2007) Involvement of transportin 2-mediated HuR import in muscle cell differentiation. *Mol Biol Cell*, **18**, 2619-2629.
161. Itoh, H., Tashima, Y., Eishi, Y. and Okeda, R. (1993) Localization of HSP90 in rat brain. *Int J Biochem*, **25**, 93-99.
162. Srikakulam, R. and Winkelmann, D.A. (2004) Chaperone-mediated folding and assembly of myosin in striated muscle. *J Cell Sci*, **117**, 641-652.
163. Doller, A., Akool el, S., Huwiler, A., Muller, R., Radeke, H.H., Pfeilschifter, J. and Eberhardt, W. (2008) Posttranslational modification of the AU-rich element binding protein HuR by protein kinase Cdelta elicits angiotensin II-induced stabilization and nuclear export of cyclooxygenase 2 mRNA. *Mol Cell Biol*, **28**, 2608-2625.
164. Mazan-Mamczarz, K., Galban, S., Lopez de Silanes, I., Martindale, J.L., Atasoy, U., Keene, J.D. and Gorospe, M. (2003) RNA-binding protein HuR enhances p53 translation in response to ultraviolet light irradiation. *Proc Natl Acad Sci U S A*, **100**, 8354-8359.
165. Dormoy-Raclet, V., Cammas, A., Celona, B., Lian, X.J., van der Giessen, K., Zivojnovic, M., Brunelli, S., Riuzzi, F., Sorci, G., Wilhelm, B.T. *et al.* (2013) HuR and miR-1192 regulate myogenesis by modulating the translation of HMGB1 mRNA. *Nat Commun*, **4**, 2388.
166. Kim, H.H., Kuwano, Y., Srikantan, S., Lee, E.K., Martindale, J.L. and Gorospe, M. (2009) HuR recruits let-7/RISC to repress c-Myc expression. *Genes Dev*, **23**, 1743-1748.
167. Abdelmohsen, K., Panda, A.C., Kang, M.J., Guo, R., Kim, J., Grammatikakis, I., Yoon, J.H., Dudekula, D.B., Noh, J.H., Yang, X. *et al.* (2014) 7SL RNA represses p53 translation by competing with HuR. *Nucleic Acids Res*, **42**, 10099-10111.
168. Uren, P.J., Burns, S.C., Ruan, J., Singh, K.K., Smith, A.D. and Penalva, L.O. (2011) Genomic analyses of the RNA-binding protein Hu antigen R (HuR) identify a complex network of target genes and novel characteristics of its binding sites. *J Biol Chem*, **286**, 37063-37066.
169. Abdelmohsen, K., Panda, A.C., Munk, R., Grammatikakis, I., Dudekula, D.B., De, S., Kim, J., Noh, J.H., Kim, K.M., Martindale, J.L. *et al.* (2017) Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol*, **14**, 361-369.
170. Simionescu-Bankston, A., Leoni, G., Wang, Y., Pham, P.P., Ramalingam, A., DuHadaway, J.B., Faundez, V., Nusrat, A., Prendergast, G.C. and Pavlath, G.K. (2013) The N-BAR domain protein, Bin3, regulates Rac1- and Cdc42-dependent processes in myogenesis. *Dev Biol*, **382**, 160-171.
171. Gallouzi, I.E., Brennan, C.M., Stenberg, M.G., Swanson, M.S., Eversole, A., Maizels, N. and Steitz, J.A. (2000) HuR binding to cytoplasmic mRNA is perturbed by heat shock. *Proc Natl Acad Sci U S A*, **97**, 3073-3078.
172. Horsley, V., Jansen, K.M., Mills, S.T. and Pavlath, G.K. (2003) IL-4 acts as a myoblast recruitment factor during mammalian muscle growth. *Cell*, **113**, 483-494.

173. Conrad, N.K., Mili, S., Marshall, E.L., Shu, M.D. and Steitz, J.A. (2006) Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. *Mol Cell*, **24**, 943-953.
174. Anantharaman, A., Gholamalamdari, O., Khan, A., Yoon, J.H., Jantsch, M.F., Hartner, J.C., Gorospe, M., Prasanth, S.G. and Prasanth, K.V. (2017) RNA-editing enzymes ADAR1 and ADAR2 coordinately regulate the editing and expression of Ctn RNA. *FEBS Lett*, **591**, 2890-2904.
175. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res*, **12**, 996-1006.
176. Mouse Genome Sequencing, C., Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M. *et al.* (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature*, **420**, 520-562.
177. Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S. *et al.* (2009) A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol*, **10**, R42.
178. Lau, A.G., Irier, H.A., Gu, J., Tian, D., Ku, L., Liu, G., Xia, M., Fritsch, B., Zheng, J.Q., Dingledine, R. *et al.* (2010) Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). *Proc Natl Acad Sci U S A*, **107**, 15945-15950.
179. Wigington, C.P., Jung, J., Rye, E.A., Belauret, S.L., Philpot, A.M., Feng, Y., Santangelo, P.J. and Corbett, A.H. (2015) Post-transcriptional regulation of programmed cell death 4 (PDCD4) mRNA by the RNA-binding proteins human antigen R (HuR) and T-cell intracellular antigen 1 (TIA1). *J Biol Chem*, **290**, 3468-3487.
180. Ule, J., Jensen, K., Mele, A. and Darnell, R.B. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*, **37**, 376-386.
181. Wigington, C.P., Morris, K.J., Newman, L.E. and Corbett, A.H. (2016) The Polyadenosine RNA-binding Protein, Zinc Finger Cys3His Protein 14 (ZC3H14), Regulates the Pre-mRNA Processing of a Key ATP Synthase Subunit mRNA. *J Biol Chem*, **291**, 22442-22459.
182. Banerjee, A., Vest, K.E., Pavlath, G.K. and Corbett, A.H. (2017) Nuclear poly(A) binding protein 1 (PABPN1) and Matrin3 interact in muscle cells and regulate RNA processing. *Nucleic Acids Res*, **45**, 10706-10725.
183. Phillips, B.L., Banerjee, A., Sanchez, B.J., di Marco, S., Gallouzi, I.E., Pavlath, G.K., Corbett, A.H. (2018), *In revision*.
184. Ha, M. and Kim, V.N. (2014) Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol*, **15**, 509-524.
185. Paziewska, A., Mikula, M., Dabrowska, M., Kulecka, M., Goryca, K., Antoniewicz, A., Dobruch, J., Borowka, A., Rutkowski, P. and Ostrowski, J. (2018) Candidate diagnostic miRNAs that can detect cancer in prostate biopsy. *Prostate*, **78**, 178-185.
186. Wang, L., Zhu, M.J., Ren, A.M., Wu, H.F., Han, W.M., Tan, R.Y. and Tu, R.Q. (2014) A ten-microRNA signature identified from a genome-wide microRNA expression profiling in human epithelial ovarian cancer. *PLoS One*, **9**, e96472.

187. Paraskevi, A., Theodoropoulos, G., Papaconstantinou, I., Mantzaris, G., Nikiteas, N. and Gazouli, M. (2012) Circulating MicroRNA in inflammatory bowel disease. *J Crohns Colitis*, **6**, 900-904.
188. Epis, M.R., Giles, K.M., Barker, A., Kendrick, T.S. and Leedman, P.J. (2009) miR-331-3p regulates ERBB-2 expression and androgen receptor signaling in prostate cancer. *J Biol Chem*, **284**, 24696-24704.
189. Hermansen, S.K., Sorensen, M.D., Hansen, A., Knudsen, S., Alvarado, A.G., Lathia, J.D. and Kristensen, B.W. (2017) A 4-miRNA signature to predict survival in glioblastomas. *PLoS One*, **12**, e0188090.
190. Chang, R.M., Yang, H., Fang, F., Xu, J.F. and Yang, L.Y. (2014) MicroRNA-331-3p promotes proliferation and metastasis of hepatocellular carcinoma by targeting PH domain and leucine-rich repeat protein phosphatase. *Hepatology*, **60**, 1251-1263.
191. Ayaz, L., Gorur, A., Yaroglu, H.Y., Ozcan, C. and Tamer, L. (2013) Differential expression of microRNAs in plasma of patients with laryngeal squamous cell carcinoma: potential early-detection markers for laryngeal squamous cell carcinoma. *J Cancer Res Clin Oncol*, **139**, 1499-1506.
192. Rehmsmeier, M., Steffen, P., Hochsmann, M. and Giegerich, R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA*, **10**, 1507-1517.
193. Cardinali, B., Castellani, L., Fasanaro, P., Basso, A., Alema, S., Martelli, F. and Falcone, G. (2009) Microrna-221 and microrna-222 modulate differentiation and maturation of skeletal muscle cells. *PLoS One*, **4**, e7607.
194. Dey, B.K., Gagan, J. and Dutta, A. (2011) miR-206 and -486 induce myoblast differentiation by downregulating Pax7. *Mol Cell Biol*, **31**, 203-214.
195. Banerjee, A., Vest, K.E., Phillips, B.L., Deng, Q., Seyfried, N.T., Pavlath, G.K., Corbett, A.H. . (2018), *In preparation*.