
DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


TRIM71 Binds to CRD Complex Members and Is Capable of Positive and Negative Regulation of Target RNAs

Citation

Foster, Daniel Jacob. 2019. TRIM71 Binds to CRD Complex Members and Is Capable of Positive and Negative Regulation of Target RNAs. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:42013145

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

TRIM71 Binds to CRD Complex Members and is Capable of Positive and Negative Regulation of Target RNAs

A dissertation presented

by

Daniel Jacob Foster

to

The Department of Biological and Biomedical Sciences

in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

July 2019

© 2019 Daniel Jacob Foster

All rights reserved

Dissertation Advisor: Frank J. Slack Daniel Jacob Foster

TRIM71 Binds to CRD Complex Members and is Capable of Positive and Negative Regulation of Target RNAs

Abstract

TRIM71 is an important RNA binding protein in development and disease. It is expressed early in development and plays a role in promoting early cell fates and preventing premature differentiation of cells. TRIM71 was discovered in *C. elegans* as a target of the *let-7* microRNA and it is conserved in humans. Its expression during development is largely controlled post-transcriptionally by *let-7* and it is downstream of the LIN28/*let-7* axis.

TRIM71 has been described as inhibiting several target mRNAs such as *CDKN1A* and *EGR1*. To better understand the role of TRIM71 in cells and the RNAs it binds, we used crosslinking IP (CLIP) followed by RNA sequencing. We discovered a number of developmentally and disease relevant RNAs that were bound by TRIM71 such as *LIN28B*, *MYC*, *MDM2,* and the *MBNL* and *NFI* families.

To understand how TRIM71 regulated RNAs in a cell we analyzed proteins that co-precipitated with TRIM71. TRIM71 was able to co-precipitate members of the CRD complex which is known to bind to and positively regulate RNAs such as *MYC*. Domain deletion analysis of TRIM71 showed that the coiled-coil domain was necessary for the interaction with DHX9, but we could not narrow down the domains interacting with the

core CRD subunit, IMP1. However, the RNA-binding NHL domain of TRIM71 did not alone bind the CRD complex members.

To determine the intrinsic RNA binding abilities of TRIM71 we depleted *IMP1* and showed that TRIM71 was able to bind RNA more promiscuously. Additionally, we performed a CLIP-Seq of the NHL domain alone and similarly found more promiscuous binding compared to wild-type TRIM71. We concluded that the CRD complex binds to TRIM71 and promotes specificity of RNA binding.

This new work shows that TRIM71 can positively regulate expression of genes with key roles in oncogenesis. We therefore analyzed the importance of TRIM71 in cancer, finding that its expression correlates with *LIN28A/B* in several cancer types. One of these, hepatocellular carcinoma, showed decreased cell proliferation and increased senescence when *TRIM71* was depleted. This work shows that TRIM71 and LIN28A/B, in addition to being linked in development, can also be linked in oncogenesis.

Acknowledgements

This work would not be possible without the help, support, and mentorship of Frank Slack and the rest of the Slack Lab. Members of the lab have made for an incredible environment and I could not have wished for anything better. I have developed what I hope will be lifelong friendships. I would like to thank Frank and the Slack Lab.

I appreciate all of the support I've received from friends and colleagues within the Biological and Biomedical Sciences program, HPREP, and the Harvard community. The community has been an incredible place to grow as a scientist. It has been wonderful to explore Boston and the Northeast with my newfound friends.

Most importantly, I would like to thank my family. My mom has been an incredible support through college and graduate school. My brother defines what it means to be a truly great role model. Lastly, I will remember my graduate experience most fondly because I met Thao. She has supported me through everything and has been a constant source of strength, even after I injured my hand which forced me to write this thesis one-handed. She has pushed and encouraged me in innumerable ways.

Thank you,

Daniel

Table of Contents

Abstract ..... **iii**

Acknowledgements ..... **v**

Table of Contents ..... **vi**

Chapter 1 – Introduction **1**

MicroRNAs and let-7 ..... *1*

LIN28/let-7 axis *4*

LIN28 as an mRNA binding protein *6*

TRIM71 (LIN-41) in development *7*

TRIM71 in stem cells *9*

TRIM71 in cancer *10*

TRIM-NHL family of proteins *12*

TRIM71 target RNAs *23*

The role of IMP1 (IGF2BP1) as an RNA binding protein *26*

RNA binding protein characteristics *28*

Unanswered questions *28*

Chapter 2 – TRIM71 interacts with CRD complex members and binds to overlapping and distinct RNAs **30**

Abstract *30*

TRIM71 binds to CRD complex members *32*

TRIM71 domains responsible for interaction with the CRD complex members *36*

TRIM71 crosslinking IP sequencing shows binding to a large number of RNAs *38*

Gene Ontology analysis of RNAs from the TRIM71 CLIP *44*

TRIM71 binds to known and novel RNAs *45*

TRIM71 regulation of target mRNAs *58*

IMP1 helps specify where TRIM71 binds *67*

RNA binding of TRIM71 domain deletion mutants *68*

Summary *83*

Chapter 3 – *TRIM71* expression correlates with *LIN28A*/*B* in cancer **86**

Abstract *86*

TRIM71 expression analysis and correlation with LIN28, CRD complex members, and RNA targets....87
The effect of TRIM71 expression on patient outcomes and growth of cell lines....95
TRIM71 in liver cancer cell lines....99
TRIM71 in neuroblastoma cell lines....105
Rescuing TRIM71 phenotypes in liver cancer cell lines....107
Summary....109

Chapter 4 – Discussion....110
Model for understanding TRIM71 as a member of the RNA-binding CRD complex....110
TRIM71 is capable of positive or negative regulation of target RNAs....112
TRIM71 NHL mutation in patients....114
TRIM71 in cancer....116
Future perspectives....117

Chapter 5 - Materials and Methods....119
Cell Culture....119
RNA extraction, Reverse Transcription, and Quantitative PCR....119
Antibodies....120
Western Blotting....120
siRNA transfections....121
MTT Assay....122
BrdU Assay....122
Senescence staining....123
Cross-linking IP and qRT-PCR....123
CLIP-sequencing....125
RNA IP....126
Plasmids and cloning....127

Bibliography....129

Chapter 1 – Introduction

MicroRNAs and *let-7*

MicroRNAs are small, non-coding RNAs that play a regulatory role in a vast array of biological processes including development, cell fate specification, and disease. These RNAs bind to the 3' untranslated region (3'UTR) of messenger RNAs (mRNAs) with imperfect complementarity and recruit the RNA induced silencing complex (RISC) to the transcript blocking translation (Stefani and Slack, 2008; Winter et al., 2009).

Since their discovery in the early 1990's in the nematode, *Caenorhabditis elegans*, many thousands of miRNAs have been annotated in most animal phyla as well as plants and viruses (Stefani and Slack, 2008). The first two miRNAs discovered, *lin-4* and *let-7*, were found to play a role during development of *C. elegans* (Lee et al., 1993; Reinhart et al., 2000).

In nematodes, mutation or deletion of *lin-4* or *let-7* causes mis-expression of target mRNAs leading to defects in timing of developmental events such as cell division or differentiation (Slack et al., 1999). Hence these genes have become known as heterochronic genes (Abrahante et al., 1998; Johnson et al., 2003). Although *lin-4* was not known to be conserved in humans at the time of discovery, *let-7* was, and so was its first discovered mRNA target, *lin-41* (Pasquinelli et al., 2000). Both *let*-7 and *lin-41* are

conserved in humans, though *lin-41* is named *TRIM71*. Similar to nematodes, human *let-7* expression increases during development as cells differentiate into various tissue types (Figure 1.1A). Characteristic of the negative regulatory role that miRNAs play, TRIM71 (LIN-41) decreases in expression as *let-7* expression goes up (Lancman et al., 2005; Lin et al., 2007; Schulman et al., 2005; Vella et al., 2004). In *let-7* mutant *C. elegans*, hypodermal seam cells remain undifferentiated and undergo additional rounds of cell division indicating that *let-7* plays an important role in inhibiting early cell fates and promoting differentiation (Großhans et al., 2005; Slack et al., 1999).

Because many *let*-7 targets are conserved from Nematodes to humans, some of the early research compared uncontrolled growth phenotypes in *C. elegans* to oncogenic phenotypes in humans (Büssing et al., 2008). The first such example was *let-60*, which has human homologs known as the *RAS* family members, *N*-, *H*-, and *KRAS*. In *C. elegans*, *let-60* is important in the development of the vulva and gain of function *let-60* worms have multiple pseudovulvae indicating that development was not properly regulated. Overexpression of *let-7* family member, *miR-84*, suppressed the pseudovulvae phenotype (Johnson et al., 2005). In humans, it was shown that *let-7* expression was lower in lung cancer compared adjacent tissue and this could explain the increased expression of *RAS* in the tumor tissue (Johnson et al., 2007). Later studies showed that loss of *let-7* aided in oncogenic transformation of cells due to losing regulation of other oncogenes such as *HMGA2* and *EZH2* (Mayr et al., 2007). Together this strongly pointed to *let-7* as a tumor suppressing miRNA. Now we know that *let-7* plays a role in many cancer types, including lung, breast, ovarian, colon, glioma, and liver cancers

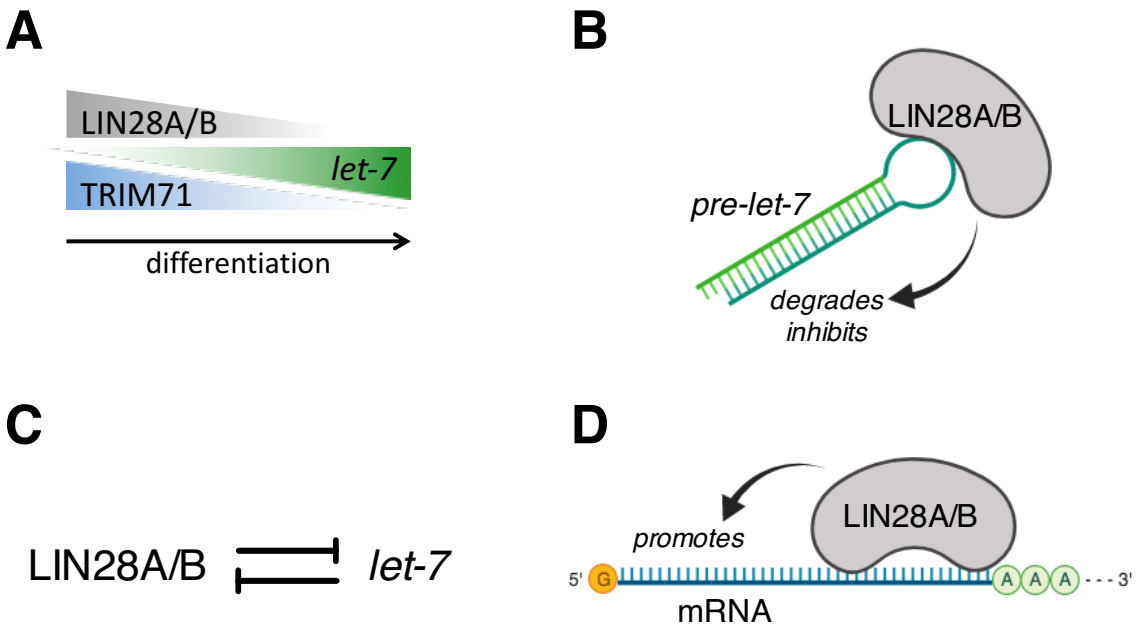


Figure 1.1. LIN28/*let-7* axis and other LIN28 mechanisms.
(A) LIN28A/B bind to the terminal stem-loop of the precursor-*let-7* microRNA resulting in both inhibition of *let-7* processing and degradation of the RNA. **(B)** *LIN28A/B* and *TRIM71* are both highly exressed early in development. As cells differentiate, LIN28 levels go down allowing for an increase in mature *let-7* followed by decreases in TRIM71 due to post-transcriptional degradation and inhibition by *let-7*. **(C)** LIN28 inihibits *let-7* processing and *let-7* is cpable of inhibiting *LIN28* post-transcriptionally. This results in a bistable switch that allows for a rapid increase in *let-7* levels during differentiation and prevents re-expression of *LIN28*. **(D)** In addition to regulating microRNA *let-7* processing, LIN28 is capable of binding mRNAs and promoting stability/translation.

(Rupaimoole and Slack, 2017; Thornton and Gregory, 2012). Overall this gives us insight as to how genes involved in development can become dysregulated, causing oncogenic phenotypes.

Let-7 actually represents a family of miRNAs with similar sequences (Roush and Slack, 2008). The human *let-7* family has 10 members and is composed of *let-7a* though *let-7i* as well as miR-98. All of these miRNAs share a similar seed sequence but differ in the rest of the sequence. The members of the *let-7* family are produced from 13 different precursor transcripts and are located throughout the human genome in 8 different loci. The expression patterns of the different family members are similar to each other and likely represent gene duplications through evolution due to the vital role they play (Roush and Slack, 2008).

LIN28/*let-7* axis

Axis in development

In organisms from C. elegans to humans, *let-7* expression is carefully regulated post-transcriptionally throughout development, most notably by the RNA binding protein LIN28. Humans have two LIN28 isoforms, LIN28A, often referred to just as LIN28, and LIN28B. The two isoforms are similar except that LIN28B contains a nuclear localization sequence, influencing its distribution within the cell (Thornton et al., 2012; Thornton and Gregory, 2012; Viswanathan and Daley, 2010; Viswanathan et al., 2008).

LIN28A/B inhibit miRNA processing by binding to the terminal end of stem-loops present in primary and precursor miRNAs (Figure 1.1B) (Nam et al., 2011). Binding of LIN28A/B to the GGAG binding motif of the *let-7* precursor disrupts the double-stranded

stem-loop to prevent further processing into a mature miRNA (Wilbert et al., 2012). LIN28A is also able to recruit binding of a Terminal Uridylyl Transferase, TUT4, which adds a 3' uridine tail to further suppress processing (Hagan et al., 2009; Heo et al., 2008; Lehrbach et al., 2009; Piskounova et al., 2011; Thornton et al., 2012).

As a negative regulator of *let-7*, *LIN28* is more highly expressed early in development, to limit the amount of mature *let-7* present in cells (Figure 1.1A). Due to the important role *let-7* plays in promoting differentiation, inhibiting its action is necessary to induce pluripotency. Therefore, LIN28 can be used as a pluripotency factor along with key embryonic stem cell (ESC) factors, OCT4, SOX2, and NANOG (Viswanathan and Daley, 2010; Yu et al., 2007).

let-7 is able to bind and repress *LIN28* as an mRNA target, so the two form a bistable switch when both are expressed in early cell fates (Figure 1.1C) (Thornton and Gregory, 2012). LIN28 regulates other miRNAs besides *let-7* and is also able to bind mRNAs containing the GGAG binding motif on stem-loop structures (Figure 1.1D) (Wilbert et al., 2012). Because it can regulate *let-7* as well as many mRNAs, the developmental and disease effects of LIN28 expression can be classified as either *let-7* dependent or *let-7* independent. However, this dichotomy is not entirely accurate, as it is possible for a single gene (e.g. *IGF2BP1*) to be regulated at the mRNA level both positively by LIN28 and negatively by *let-7*. Nevertheless, it is often important to consider both *let-7* dependent and *let-7* independent pathways as they can have separate effects or converge on a common goal. In the case of *IGF2BP1*, expression is directly promoted by LIN28 binding to the mRNA and degradation is prevented by LIN28 inhibiting *let-7*.

LIN28 in cancer

Consistent with its role in promoting pluripotency, *LIN28* is an oncogene in many cancers. In over 10 different cancer types, high *LIN28A/B* expression is correlated and predictive of worse prognosis and more severe cellular phenotypes (Jiang and Baltimore, 2016; Murray et al., 2013; Nguyen et al., 2014; Tu et al., 2015; Urbach et al., 2014). Likewise, *LIN28* expression is anticorrelated with *let-7* expression. In cancer, increased *LIN28* expression is rarely due to amplification (Viswanathan et al., 2009). Instead, *LIN28* is more commonly overexpressed due to transcriptional activation or locus hypomethylation. SOX2 is likely promoting *LIN28* expression in the developmental context (Buganim et al., 2012; Shyh-Chang and Daley, 2013) and MYC and NF-kB could be regulating expression in the oncogenic context (Viswanathan et al., 2009). LIN28 is able to act through both *let-7* dependent and independent mechanisms in development and disease.

LIN28 as an mRNA binding protein

Given the small RNA binding motif of LIN28, it does not bind with great specificity. Studies looking at all bound RNAs using CLIP-seq have found discrete binding sites in around 25% of all human transcripts on the low end to nearly all expressed transcripts on the high end (Cho et al., 2012; Graf et al., 2013; Wilbert et al., 2012). In contrast to its role in inhibiting miRNA biogenesis, as an mRNA binding protein, LIN28 acts mostly as a positive regulator, promoting translation (Graf et al., 2013). LIN28A interacts directly and indirectly with translation initiation factors, poly(A)-binding protein (PABP), and ribosomal

components to regulate translation. Interestingly, the regulation of LIN28 on mRNAs seems to be location dependent, as targets bound within the CDS are more strongly regulated than those bound in the 3'UTR (Graf et al., 2013).

IGF2BP1 (*IMP1*) as well as family members *IGF2BP2/3* (*IMP2/3*) are important targets of LIN28B, especially in the context of development and cancer. *IMP1* and *IMP3* are highly expressed in a *LIN28B* overexpression mouse model of liver cancer and all are known to be direct mRNA targets of LIN28B (Nguyen et al., 2014). The *IMP* family are RNABPs themselves and are known to promote cancer progression through targeting of mRNAs such as *MYC* and *HMGA1* (Conway et al., 2016; Degrauwe et al., 2016a; Degrauwe et al., 2016b; Müller et al., 2018).

TRIM71 (LIN-41) in development

In *C. elegans*, LIN-41 is involved with timing of development events and is normally expressed in early stages of development (Figure 1.1). Mutant *lin-41 C. elegans* exhibit precocious phenotypes where developmental events occur too early. Formation of alae is a common marker of later development and it normally forms as worms are transitioning from the L4 larval stage to adulthood. Alae is a structure formed on the outside cuticle of the worm, secreted by hypodermal seam cells. In *lin-41* mutant worms, alae forms earlier than expected, as worms transition from the L3 to L4 larval stage (Slack et al., 1999). *lin-41* is also expressed in the gonads of worms and mutation in *lin-41* decreases the number of progeny within each brood.

TRIM71 (*trim71*) is an essential gene in zebrafish development. Around 35-50% of embryos that are injected with *trim71* small interfering RNAs (siRNAs) exhibit severe

defects and begin to die around 24-48 hours after injection (Lin et al., 2007). Around half of the remaining embryos exhibited abnormal phenotype of varying degrees of severity.

In mice, *Trim71* is expressed ubiquitously in early embryogenesis, around day E9.0 of development. By day E10.5, *Trim71* decreases in expression and is localized to limb buds and certain areas of the brain. Within the limb buds, expression is further localized to the posterior and distal ends (Schulman et al., 2005). Meanwhile, *let-7* is more localized to the anterior limb buds (Mansfield et al., 2004), indicating that *let-7* can control the temporal and spatial expression patterns of *Trim71* in developing embryos.

Mice that have *Trim71* deleted exhibit 100% embryonic lethality by day E16.5, indicating that it is essential in development. These mice have a severe neural tube closure defect (Schulman et al., 2008). Confocal microscopy of mouse brain sections showed that *Trim71* was localized to neural progenitor cells which make up the neural tube (Chen et al., 2012). Furthermore, loss of *Trim71* decreased the proliferation of these cells and caused premature differentiation. *Trim71* knockout mES cells showed premature expression of proneural genes that promote neural differentiation (Mitschka et al., 2015).

The canal running through the neural tube eventually becomes the ventricles of the brain. In humans, *TRIM71* mutations have been linked to congenital hydrocephalus, a disease in which excess cerebrospinal fluid builds up in the ventricles causing distension and leading to high mortality and severe developmental defects (Furey et al., 2018). Although the exact targets of TRIM71 are unknown, its importance in development of species from *C. elegans* to humans is clear.

TRIM71 in stem cells

Inhibition of *let-7* during reprogramming of fibroblasts is important to increase yield of ES-like iPS cells. *TRIM71* was the main *let-7* target affected during reprogramming, and overexpression of *TRIM71* played a similar role as *let-7* inhibition in facilitating iPS cell formation (Worringer et al., 2014). *TRIM71* in combination with *OCT4*, *SOX2*, and *KLF4* (OSK) greatly facilitated reprogramming compared to OSK alone. Interestingly, when OSK was expressed along with *MYC*, *TRIM71* became dispensable, indicating that they play redundant roles, possibly in downstream gene regulation. It has been shown that the *TRIM71* promoter has MYC binding sites and is sensitive to MYC levels (Chen et al., 2013) so overexpression of *MYC* could be driving *TRIM71* levels in these cells and account for the redundancy. An alternative explanation is that TRIM71 could be regulating *MYC* levels in these cells. We show here that TRIM71 is capable of binding to and regulating the *MYC* mRNA.

TRIM71 is highly expressed in both human and mouse stem cells (Chang et al., 2012; Mitschka et al., 2015; Rand et al., 2018; Worringer et al., 2014). Knockout of *Trim71* in mouse embryonic stem (ES) cells increases expression of genes normally seen in neural stem cells and neurons (Mitschka et al., 2015). Gene ontology (GO) analysis further supported the idea that *Trim71* knockout cells are primed for neural differentiation. Enriched GO categories include “Nervous system development” and “Neurogenesis, neural cell morphogenesis involved in differentiation, and axogenesis”. Additionally, it was shown that in an embryoid body assay followed by retinoic acid stimulation, *Trim71* knockout cells had much higher expression of neuroectodermal markers. *Trim71* knockout also increases the expression of brain-associated miRNAs. These studies in mES cells as well as studies in vivo show that TRIM71 is crucial in the regulation of neural

development and disruption of *TRIM71* dysregulates the timing of neural differentiation (Chen et al., 2012; Lin et al., 2007; Mitschka et al., 2015).

In addition to repressing the neural development program, TRIM71 plays a role in promoting proliferation of mouse ES cells and human IPS cells by post-transcriptionally repressing *CDKN1A*, the mRNA encoding p21 (Chang et al., 2012; Rand et al., 2018). CDKN1A(p21) is a cyclin-dependent kinase inhibitor that negatively regulates the G1-S transition phase of the cell cycle by binding to an inhibiting CDK2 and CDK4. In the presence of TRIM71, *CDKN1A* translation is inhibited thereby increasing the activities of CDK2 and CDK4, resulting in increased proliferation of cells. One effect of this is increased phosphorylation and decreased activity of RB1, a negative regulator of the cell cycle (Rand et al., 2018).

TRIM71 in cancer

Given the role TRIM71 plays in development and stem cells in promoting cell proliferation and preventing neuronal differentiation, we have an interest in understanding TRIM71 in the context of cancer. *LIN28* and *let-7* have provided examples of developmental heterochronic genes that also play important roles in disease.

TRIM71 as an oncogene

There are conflicting studies indicating that TRIM71 could play an oncogenic role or a tumor suppressive role. *TRIM71* is highly expressed in hepatocellular carcinoma (HCC) and knockdown of *TRIM71* in HCC cell lines decreases proliferation and tumor

formation in xenografts which would point towards an oncogenic role for *TRIM71* (Chen et al., 2013).

In non-small cell lung cancer (NSCLC), relative to liver cancer, *TRIM71* is not very highly expressed. However, the expression of *TRIM71* still correlates with clinical and pathological characteristics such as tumor size, differentiation, metastasis, and cancer stage in the dataset analyzed (Ren et al., 2017). In NSCLC cell lines, knockdown of *TRIM71* slowed cell growth and proliferation. Here, TRIM71 played a role in activating the NF-κB pathway by decreasing levels of IκBα, an NF-κB inhibitor. The NF-κB pathway promotes cell proliferation and survival so if TRIM71 promotes the activity of this pathway it would support the idea that TRIM71 is an oncogene. The authors hypothesized that the TRIM71 RING domain was responsible for ubiquitinating IκBα leading to its degradation, however, they were not able to show direct interaction of TRIM71 with IκBα.

It has also been suggested that p53 could be a ubiquitination target of TRIM71 (Nguyen et al., 2017). However, overexpression of *TRIM71* caused miniscule changes in p53 ubiquitination, if any. It was shown that RNA-binding NHL domain, but not the E3 ubiquitin ligase domain, was necessary and sufficient for TRIM71 to co-IP with p53. While it is possible that the NHL domain could have a role in protein binding, x-ray crystallography studies suggest otherwise (Kumari et al., 2018).

TRIM71 as a tumor suppressor

Several studies have looked at the role of TRIM71 as a tumor suppressor and all relate to its ubiquitination activity as an E3 ubiquitin ligase. The potential ubiquitination targets important in cancer include AGO2 and LIN28B (Lee et al., 2014; Rybak et al.,

2009; Yin et al., 2016). Some of these studies have failed to replicate in physiological conditions in vivo, including AGO2 ubiquitination that was initially reported in embryoid carcinoma cells (Chang et al., 2012). For LIN28B, some reports indicate that TRIM71 is inhibiting it through ubiquitination while other reports show TRIM71 cooperating with LIN28. Consequently, these studies report opposite effects on *let-7* regulation by TRIM71. Regulation of these proteins by TRIM71 has not been shown to cause any phenotypic impacts on cells.

Taken together these various studies point to several possible ubiquitination targets but the lack of reproducibility could be explained by different cell types used in the studies. The RING domain of TRIM71 is conserved from *C. elegans* to humans which would indicate there is an important role for TRIM71 as an E3 ubiquitin ligase. However, the ubiquitination targets remain elusive.

TRIM-NHL family of proteins

TRIM71 is a member of the tripartite motif (TRIM) and NHL families (Figure 1.2A). The TRIM family is composed of around 75 members that play roles in development, cancer, autophagy, immunity, etc. (Hatakeyama, 2017). The tripartite motif is defined by the presence of a RING-finger E3 ubiquitin ligase domain, one or two zinc-binding B-box domains, and a coiled-coil domain. Many but not all TRIM proteins have a functional RING domain (Hatakeyama, 2011; Hatakeyama, 2017). The B-box domains may also play a role in the ubiquitination process while the coiled coil domain is likely involved in protein binding. It has been proposed that the B-box and coiled-coil domains provide a binding site for substrates that are ubiquitinated by the RING domain (Tocchini and Ciosk, 2015).

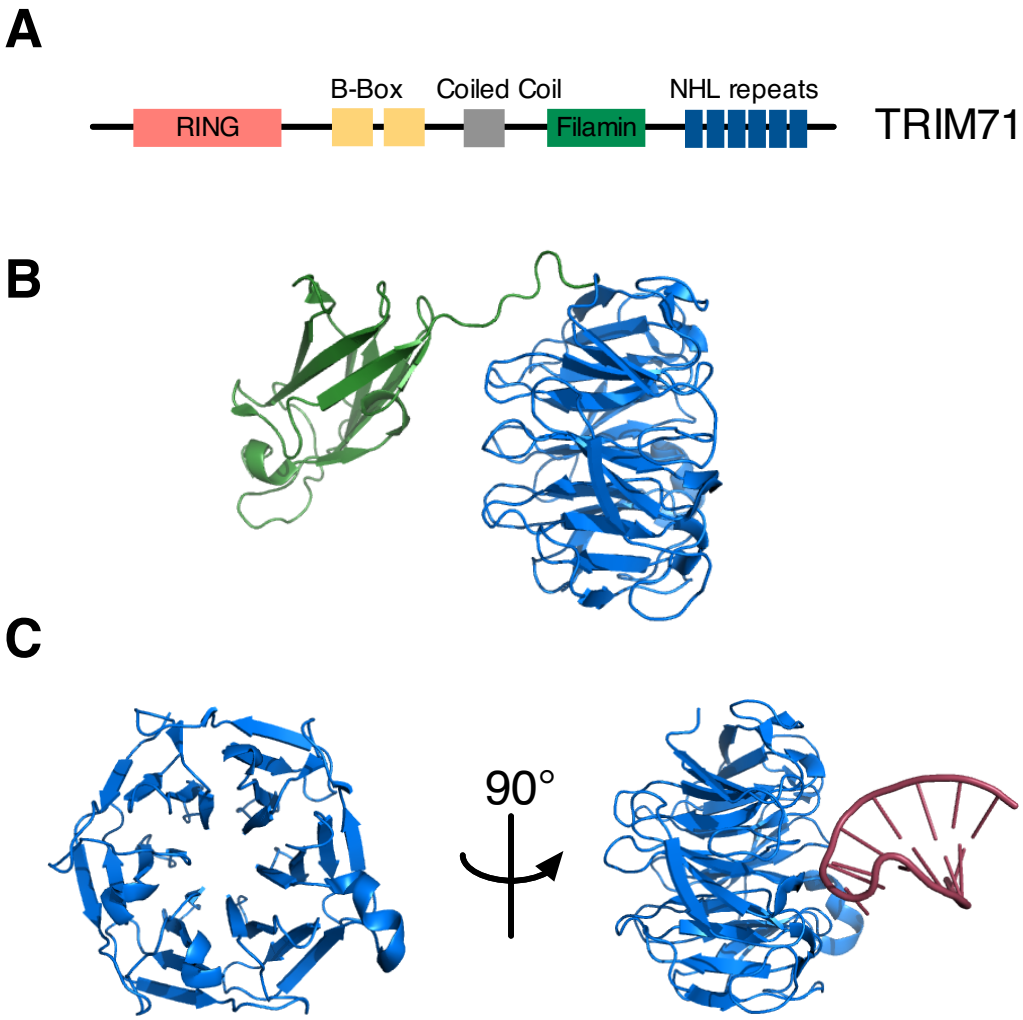


Figure 1.2. TRIM71 domains and crystal structures.
(A) Schematic of TRIM71 domains. The E3 ubiquitin-ligase RING domain is at the N-terminal end of TRIM71 whereas the RNA-binding NHL domain is at the C-terminus. **(B)** Crystal structures of the *D. rario* Trim71 filamin and NHL domains published in Kumari, *et al*., 2018 were downloaded from the Protein Data Bank (6FQL) and analyzed in PyMol. **(C, left)** The NHL domain forms a six-bladed propellar-like structure. **(C, right)** The NHL domain is shown bound to RNA. Here, a small portion of the *C. elegans mab-10* 3'UTR is shown bound to *D. rario* Trim71.

A subset of TRIM family members contains an RNA-binding NHL domain. The NHL family of proteins is named after repeats within the NCL-1, HT2A, and LIN-41 proteins in *C. elegans* (Slack and Ruvkun, 1998). The RNA binding of TRIM proteins represents an additional molecular function that is possibly independent from the ubiquitin ligase activity. The various TRIM-NHL family members have different sequence preferences for binding RNA and are involved in a variety of developmental processes (Loedige et al., 2015). In *Drosophila melanogaster*, the TRIM-NHL protein, Brat, is associated with differentiation vs self-renewal fate decisions of the neuroblast cell lineage (Tocchini and Ciosk, 2015). Another *D. melanogaster* TRIM-NHL protein, Mei-P26, is important in germline development. Trim2, Trim3, and Trim32 are important for neuronal development in mice. As a result of the developmental processes TRIM-NHL proteins play roles in, mutations and aberrations in expression have been associated with a variety of diseases and cancers (Tocchini and Ciosk, 2015).

TRIM71 Tripartite motif

As the tripartite motif is important for the ubiquitination function, most studies looking at TRIM71 in this respect have validated their results by mutating or deleting the RING domain or the entire TRIM motif (Table 1.1).

Although AGO2 was convincingly shown to be a ubiquitination target of TRIM71 in P19 cells, other studies have failed to replicate these results in other cells or under physiological conditions in vivo (Chen et al., 2012). Ago2 is was found to co-precipitate with Trim71 in mES cells but was not a target of ubiquitination (Chang et al., 2012).

Table 1.1: Proposed TRIM71 ubiquitination targets and validation studies

Ubiquitination target	Cell line	Study	Validation experiment
AGO2	P19 Embryonal carcinoma	Rybak, 2009	TRIM deletion
LIN28B	HEK293	Lee, 2014; Yin, 2016	RING mutation
p53	Mouse ES cells	Nguyen, 2017	none
IκB-α	A549 NSCLC	Ren, 2017	RING deletion

LIN28B was shown to be a ubiquitination target in HEK293 cells when *TRIM71* was overexpressed (Lee et al., 2014). However, LIN28B was also ubiquitinated in HEK293 cells even in the absence of *TRIM71*. While expression of a RING domain mutant *TRIM71* did decrease LIN28 ubiquitination relative to wild-type *TRIM71* expression, but it did not completely ablate it. *TRIM71* has very low basal expression in these cells so LIN28B may not be a physiologically relevant ubiquitination target. Lastly, LIN28B and TRIM71 play complementary roles as pro-growth factors during development and reprogramming, as they are often co-expressed, so it is unclear how this potential mode of regulation would affect cells in development. Future studies looking at an inhibitory role of TRIM71 on LIN28B should be undertaken in a physiologically relevant context.

It has been proposed that TRIM71 plays an important role in regulating p53 functions within a cell. During differentiation of mES cells with retinoic acid, levels of p53 were decreased in cells with overexpressed *TRIM71* (Nguyen et al., 2017). However, any increase in ubiquitination of p53 with *TRIM71* overexpression was slight. Unfortunately,

this study did not have validation experiments looking at p53 ubiquitination with the RING domain deleted or mutated, so any effects could be indirect.

The final proposed ubiquitination target of TRIM71 was IκB-α in the context of non-small cell lung cancer (Ren et al., 2017). Although *TRIM71* is lowly expressed in the lung, there appears to be higher expression in the cancer tissue compared to normal, and high expression correlated with decreased survival, increased metastasis, and increased tumor size in the data sets analyzed. This study showed that knockdown of *TRIM71* decreased ubiquitination of IκB-α and overexpression of full-length *TRIM71* but not a RING deletion mutant, increased ubiquitination of IκB-α.

The tripartite motif of TRIM71 has been conserved from *C. elegans* to humans and therefore could play an important role in the function of the protein. However, as shown above, the currently described targets of TRIM71 ubiquitination are unclear, and some may not be physiologically relevant. Interestingly, the worm version of TRIM71, LIN-41, may not have a functional role as an E3 ligase in *C. elegans* as the protein does not share key residues with other E3 ligases that should be required for interaction with E2 proteins (Tocchini et al., 2014). Furthermore, worms with mutations in the NHL domain have germline and somatic phenotypes that are fully rescued by expressing *lin-41* with mutations in the RING domain indicating that the ubiquitination function is not responsible for observed phenotypes (Tocchini et al., 2014).

As TRIM71 is also an RNA binding protein, it is possible that the main ubiquitination targets are other proteins that are simultaneously bound to the same RNAs as TRIM71. This is in fact the mechanism for other RNA binding proteins that also have E3 ubiquitin ligase activity such as Dzip3 and Mex3b which rely on the lncRNA HOTAIR as a scaffold to interact with substrates Ataxin-1 and Snurportin-1 (Yoon et al., 2013). However, given

that RING mutant TRIM71 can rescue *C. elegans* phenotypes and can still reprogram fibroblasts, the roles of ubiquitination remains unclear.

TRIM71 NHL domain

The NHL domain is perhaps the most interesting domain of TRIM71 because most mutations map to this region. Early studies to uncover *let-7* targets found that mutations in the NHL domain of *C. elegans lin-41* could suppress the phenotypes associated with *let-7* mutation (Slack et al., 1999). More recently, it was shown in humans that mutations in the NHL domain of *TRIM71* could lead to congenital hydrocephalus (Furey et al., 2018). This points to the NHL domain and the RNA binding properties of TRIM71 as being the most important for the developmental phenotypes seen in *C. elegans* and humans.

The NHL domain is characterized by 5 or 6 repeats that form a propeller-like structure that binds RNA (Figure 1.2B, C). It is thought that the NHL domain provides the specificity of binding for NHL proteins as NHL domain swap experiments between TRIM71 and TRIM32 alter the RNA binding ability of the proteins (Loedige et al., 2013). A chimeric version of TRIM32 with the NHL domain of TRIM71 is able to repress proposed TRIM71 targets such as *CCNE2* and *HOXA5*. However, the corresponding chimeric version of TRIM71 with the NHL domain of TRIM32 was not able to repress the proposed mRNA targets (Loedige et al., 2013).

The outside face of the NHL domain is positively charged and capable of binding the negatively charged backbone of RNA. X-ray crystallization studies show that the NHL domain interacts with about a 6-mer RNA sequence. The NHL domain and RNA interact through a combination of aromatic staking of side chains with the RNA bases, ionic

interactions between side chains and RNA phosphate backbone, and hydrogen bonds with the bases and sugars of the RNA (Kumari et al., 2018; Loedige et al., 2015).

Several different RNA recognition motifs have been proposed for TRIM71. The first is based on an *in vitro* RNAcompete experiment in which predesigned RNA oligonucleotides of varying length were incubated with purified protein. The protein was then pulled down and the RNAs were sequenced (Ray et al., 2017). When RNAcompete was performed with human TRIM71, a GUAAAGC consensus motif was found (Loedige et al., 2015). However, this sequence is not present in any of the proposed mRNA targets of TRIM71 such as *CDKN1A* or *EGR1*; so this RNAcompete experiment which is done with purified protein *in vitro*, may not predictive motif information. A meta-analysis of RNAcompete datasets, which included the previous study, later showed that TRIM71 actually has less sequence specific recognition when compared with other NHL proteins. Instead, it has stronger recognition for a structural motif within the RNA (Figure 1.3A) (Kumari et al., 2018). Interestingly, a majority of the RNABPs analyzed in the meta-analysis of 260 datasets did have sequence motif preference with TRIM71 and LIN-41 being the main exceptions to this. Instead of looking for 7-mer sequence motifs within the bound RNA oligos, Kumari *et al*. looked for 11-mer structural motifs and found that TRIM71 was preferentially binding to stem-loop structures with 3 nucleotide loops at the terminal end. The structure motif preference of LIN-41 and TRIM71 was confirmed by reporter constructs with hairpin loops in *C. elegans* and HEK293 cells. Mutating one nucleotide to break the hairpin structure prevented LIN-41/TRIM71 from binding while a complementary mutation to reform the hairpin allowed LIN-41/TRIM71 to bind.

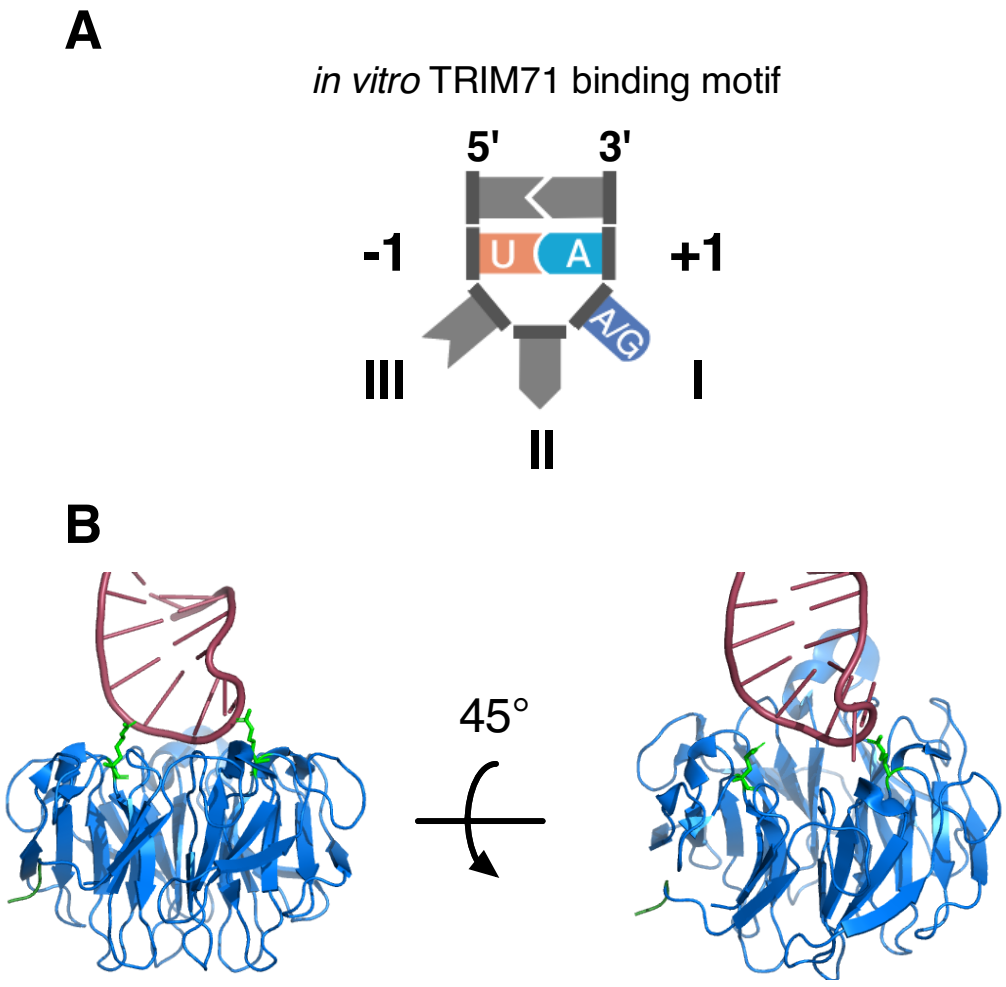


Figure 1.3. TRIM71 binding motif and NHL point mutations.
(A) Proposed TRIM71 binding motif showing structural and sequence preferences based on *in vitro* RNA binding experiments. TRIM71 preferentially binds to stem-loops with 3 nucleotides terminal loops. There is a further preference for a purine (A/G) in position I, and a U:A pair at positions -1 and +1, respectively. **(B)** Crystal structures of the *D. rario* TRIM71 NHL domain published in Kumari, *et al*., 2018 were downloaded from the Protein Data Bank (6FQL) and analyzed in PyMol. The two residues mutated in congenital hydrocephalus patients from Furey, et al., 2018, R608 and R797, are shown in green. These two residues are located within the RNA-binding pocket of the NHL domain and their side chains directly interact with the RNA.

Crystallization of the zebrafish *trim71* NHL domain with an RNA oligonucleotide showed that the NHL domain was interacting with the 3 nucleotides in the stem-loop as well as the 2 nucleotides to the 3' end of the loop structure. A total of 15 residues within the NHL domain were interacting with the 5 RNA nucleotides (Kumari et al., 2018). Notably, the NHL mutations seen in congenital hydrocephalus patients, R608H and R796H, are residues within the binding pocket that directly interact with the target RNA (Figure 1.3C) (Furey et al., 2018). Furthermore, mutations of these specific residues were shown to affect the ability of TRIM71 to regulate a reporter construct of a target RNA (Kumari et al., 2018).

The sequence motif GUAAAGC from the previous *in vitro* RNAcompete experiment does not form a hairpin loop structure based on its sequence so it's still unclear whether TRIM71 is binding preferentially to a sequence or a structural motif. As all experiments thus far have been conducted in an *in vitro* context, it is not clear whether the RNA oligonucleotides pulled down with the RNAcompete methodology are predictive of RNA sequences bound under physiological conditions.

Analysis of domain mutants – TRIM71 regulation of RNA

Current literature suggests that when TRIM71 binds to a target RNA it is repressed (Chang et al., 2012; Kumari et al., 2018; Loedige et al., 2013; Mitschka et al., 2015; Rand et al., 2018; Worringer et al., 2014). Tethering of Trim71 to the 3'UTR of a luciferase reporter construct with an MS2 fusion also shows repression (Chang et al., 2012). Table 1.2 shows what domains were required for repression of a luciferase reporter. The RING domain which is involved in E3 ubiquitin ligase activity is not required for reporter

repression. The B-box domains involved in protein binding are also not required. However, deletion of the entire TRIM motif did prevent silencing so the coiled-coil domain appears important. However, there were no deletion mutants with just the coiled-coil domain deleted to confirm this. Likewise, the filamin domain appears important but there were no single domain deletions to confirm this. Deletion of the NHL domain prevents the protein from binding to RNA, however, if the protein binds through other means such as MS2 tethering, the NHL domain is not required to repress the RNA. It appears that the minimum domains required for repressing RNAs are the coiled-coil and filamin domains. However, this was not explicitly tested and can only be inferred based on the various deletion constructs described above.

Table 1.2. TRIM71 domains important for repressing luciferase reporter

Domain	Required for luciferase reporter repression	Reference
RING	No	Chang, 2012 Loedige, 2013
B-box	No	Chang, 2012 Loedige, 2013
Coiled-coil	Yes, not tested on its own	Chang, 2012 Loedige, 2013
Filamin	Yes, not tested on its own	Chang, 2012
NHL	Yes (not necessary if protein is tethered by MS2)	Chang, 2012 Loedige, 2013

Analysis of domain mutants – TRIM71 as a pluripotency factor

As described above, TRIM71 can act as a pluripotency factor when expressed in combination with *OCT4*, *SOX2*, and *KLF4*, yet becomes redundant when *MYC* is also expressed (Worringer et al., 2014). To determine what domains of TRIM71 were important for reprogramming, Worringer *et al.* made deletions in *TRIM71* then expressed the mutants in combination with OSK and measured reprogramming efficiency (Table 1.3).

Mutations within the RING domain that prevent ubiquitin ligase activity did not decrease reprogramming efficiency however complete deletion of the RING domain did. This would suggest that ubiquitin ligase activity of TRIM71 is not required but there could be important protein-protein interactions facilitated by the RING domain that are important. Interestingly, in the presence of *MYC* expression, when *TRIM71* is typically redundant, expression of the *ΔRING-TRIM71* severely decreased reprogramming efficiency. Furthermore, high expression of *ΔRING-TRIM71* in hES cells caused a lot of cell death indicating that this mutant could be toxic. Therefore, it was difficult for Worringer *et al*. to ascribe a particular role to the RING domain during reprogramming, other than concluding that the enzymatic activity of ubiquitination was not necessary.

Deletion of the B-box, coiled-coil, or filamin domains caused a small decrease in reprogramming efficiency. These mutants also showed different localization within the cell compared to full length TRIM71. Deletion of the NHL domain as well as expression of the NHL domain on its own were not able to facilitate reprogramming. This indicates that although RNA binding ability of TRIM71 is necessary for reprogramming, it alone is not sufficient.

Table 1.3. TRIM71 domains important for reprograming fibroblasts into iPS cells.

TRIM71 domain mutants	Reprogramming efficiency with noted mutants
ΔRING	Low
RING mutations	High
ΔB-box	Medium
ΔCoiled-coil	Medium
ΔFilamin	Medium
ΔNHL	Low
NHL-only	Low

TRIM71 target RNAs

Most of the important roles of TRIM71 require binding to RNA so there has been much interest in determining what the RNA targets are (Table 1.4). The first study to look at an RNA target found that Trim71 was binding to and inhibiting the mRNA of *Cdkn1a* in mouse ES cells (Chang et al., 2012). Assuming that mRNA targets would be negatively regulated, the next study performed RNA-seq in hES cells with *TRIM71* knocked down and looked for upregulated mRNAs. They found *EGR1* as the most upregulated mRNA and further showed that TRIM71 was directly binding and inhibiting the mRNA

Table 1.4. TRIM71 target mRNAs and validation experiments.

Target mRNA	**Affects mRNA level** (if yes, cell type tested in)	**Affects luciferase reporter of 3'UTR** (if yes, cell type tested in)	**Affects protein level**
CDKN1A	mES cells ES cells	mES	yes
EGR1	mES cells ES cells	not tested	yes
CCNE2	HEK293 mES cells	HEK293 mES cells	not tested
E2F7	HEK293 mES cells	HEK293	not tested
HOXA5	HEK293	HEK293	not tested
HOXA6	HEK293	no	not tested
STAT5B	HEK293	HEK293	not tested
ZNF362	HEK293	HEK293	not tested
AIOLOS	HEK293	not tested	not tested
MYB	HEK293	no	not tested
CDH1	HEK293	HEK293	not tested
Gm2694	not tested	not tested	N/A

(Worringer et al., 2014). EGR1 is a pro-differentiation transcription factor so they concluded that the role of TRIM71 was to inhibit differentiation. Although Worringer *et al*. did observe *CDKN1A* to be a target, they concluded that it was not the main target responsible for reprogramming. However, a follow-up study from the same group did in fact determine that *CDKN1A* was an important target in reprogramming of cells (Rand et al., 2018).

Later, TRIM71 was studied in HEK293 cells when overexpressed (Loedige et al., 2013). Using mRNA arrays they found 524 downregulated mRNAs and 297 upregulated mRNAs. Next they performed RNA-IP (RIP) of FLAG-TRIM71 followed by qPCR for differentially expressed genes and found that 9 RNAs were enriched in the RIP, though only were 3 were strongly enriched. With the hypothesis that TRIM71 was only inhibiting mRNAs, they did not follow up on genes enriched in the *FLAG-TRIM71* overexpressing cells. Luciferase reporter assays with the 3'UTRs confirmed inhibition of 6 of the 9 proposed target mRNAs when *TRIM71* was overexpressed. These include *CCNE2*, *HOXA5*, *STAT5B*, *ZNF362*, *E2F7*, and *HOXA6*. Follow-up on these targets in mES cells showed that only *Ccne2* and *E2f7* were affected at the mRNA level. Luciferase reporter experiments confirmed *CCNE2* as a target. Overall, validation of targets in HEK293 cells, which do not endogenously express *TRIM71,* proved difficult.

Luciferase reporter tethering assays in both HEK293 and mES cells, where TRIM71 was tethered to a reporter mRNA, showed ~20% decrease in the mRNA level of the reporter despite seeing a 60-80% decrease in reporter activity (Loedige et al., 2013). This would suggest that the effects of TRIM71 on the mRNAs may be related to translation of the transcripts as opposed to stability.

Recently, long non-coding RNAs (lincRNAs) were evaluated as potential Trim71 targets in mES cells (Li et al., 2019). RNA IP showed that several lincRNAs were bound by Trim71 with Gm2694 being the most enriched in terms of FPKM relative expression. The effects of Trim71 on the RNA levels of Gm2694 were not tested. Rather the hypothesis was that Gm2694 acts as a sponge for Trim71 preventing it from stabilizing a protein target Shcbp1. As many lincRNAs are not conserved from mice to humans, there was no human homolog of Gm2694 to validate this interaction across species.

The role of IMP1 (IGF2BP1) as an RNA binding protein

IMP1 in development and cancer

The insulin-like growth factor 2 mRNA binding proteins (IGF2BPs) are a family of proteins expressed in embryogenesis that play a role in regulating RNAs localization, translation, and stability (Degrauwe et al., 2016b). There are three family members *IGF2BP1/2/3*, that are known by many names including *IMP1/2/3*. These proteins play a role in development of many tissues including the brain where IMP1 promotes self-renewal and prevents premature differentiation of neurons (Nishino et al., 2013). Mice with homozygous *IMP1* deletion are around 40% smaller than wild-type or heterozygous mice. Around 50% of *IMP1* -/- mice die shortly after birth within 3 days (Hansen et al., 2004).

These are oncofetal proteins that can become re-expressed in cancer and play an oncogenic role. IMP1 is one of the better characterized family members due to its role in stabilizing RNAs. In cancer, IMP1 is mainly implicated in binding to and promoting the stability of oncogenic RNAs such as *MYC*, *βTr-CP1*, *GLI1*, *MITF*, and *MDR1* (Doyle et al.,

1998; Elcheva et al., 2009; Noubissi et al., 2006; Sparanese and Lee, 2007; Weidensdorfer et al., 2009). It is a direct transcriptional target of β-catenin/TCF and is thus highly expressed in many cancers where the Wnt pathway is constitutively active (Noubissi et al., 2006; Noubissi et al., 2009).

IMP1 mechanisms

The *IMP* family plays a role in regulating RNA localization and is capable of shuttling RNAs into mRNA ribonucleoprotein (mRNP) granules. mRNP granules are self-assembled structures within a cell that contain mRNA not currently translating, non-coding RNAs, ribosomal components, and other RNA regulatory proteins. The functions of the granules vary based on their components and they can play a role is stabilizing or degrading RNA, transporting RNA, etc. (Buchan, 2014). Some reports refer to IMP1 associated granules as P-bodies while others report them as being distinct from P-bodies (Degrauwe et al., 2016b). However, the role seems to be clear that the RNAs in these granules are protected from degradation by nucleases.

IMP1 binds a large number of RNAs is able to recognize RNA N6-methyladenosine (m6A) modifications which could help specify where and to which RNAs IMP1 binds (Conway et al., 2016; Huang et al., 2018). IMP1 has been shown to interact with the RNAs of many *let-7* targets such as *LIN28B*, *HMGA2*, and *MYC*. For some RNA targets, it stabilizes them by blocking *let-7* from binding. For others, such as *MYC*, it binds to a highly structured, AU-rich element containing rare codons within the coding region called the coding region instability determinant (CRD). This region is prone to ribosome stalling and in the absence of IMP1 the RNA is quickly degraded (Doyle et al., 1998; Lemm and

Ross, 2002; Noubissi et al., 2006; Sparanese and Lee, 2007; Weidensdorfer et al., 2009). IMP1 binds to the CRD element with a group of proteins collectively called the CRD complex. Most of the various complex members have their own RNA binding motifs and play a complementary role in stabilizing the RNA by recruiting translation factors and by other unknown mechanisms.

RNA binding protein characteristics

RNA binding domains typically recognize a 4-6 nucleotide sequence of RNA. Some RNA binding proteins have multiple RNA binding domains which can allow for recognition of longer stretches of RNA or combine multiple weak interactions to achieve higher affinity and specificity (Mackereth and Sattler, 2012; Müller-McNicoll and Neugebauer, 2013). This is the case for LIN28 which contains a cold shock domain as well as two zinc finger domains, all of which can bind to RNA (Nam et al., 2011). Presumably the same multidomain interactions can be observed between different RNA binding proteins. However, it is difficult to study the structure of multidomain proteins and their interactions with other multidomain proteins due to the difficulty in crystallization and the potentially different binding characteristics between crystalized protein and protein in solution (Mackereth and Sattler, 2012).

Unanswered questions

For TRIM71, the biggest questions that remain to be answered are 1) what are the target RNAs of TRIM71, 2) by what mechanism does TRIM71 regulate these RNAs, and 3) what role if any does TRIM71 ubiquitination have within a cell. Although there are some insights into these questions, there is still much to discover There has been no unbiased

characterization of TRIM71 bound RNAs. The current methods of target finding have revolved around finding differentially expressed RNAs through RNA-seq or arrays with *TRIM71* depletion or overexpression in cells, followed by characterizing whether TRIM71 binds those RNAs. Since the mechanisms of regulation by TRIM71 are unknown, it is not clear that TRIM71 binding to RNA will necessarily affect the levels of that RNA. There could be additional targets yet to be elucidated which are regulated at the translational level, for which RNA-seq would not necessarily detect. In this thesis we hope to add to the understanding of what RNAs are bound through the analysis of CLIP-seq and how TRIM71 might regulate them by cooperating with CRD complex members.

Chapter 2 – TRIM71 interacts with CRD complex members and binds to overlapping and distinct RNAs

Authors: Daniel J. Foster, Hao-Ming Chang, Jeffrey R. Haswell

Contributions: Hao-Ming Chang performed the IP-MS of Trim71 in mouse ES cells. Jeffrey Haswell performed two replicates of the TRIM71 RIP with *IMP1* depleted. Daniel Foster performed all other experiments.

Abstract

Trim71 co-precipitates with all members of the CRD complex in mouse ES cells. In human cells we confirmed binding to CRD complex members IMP1 and DHX9 and found that the coiled-coil domain mediated the TRIM71-DHX9 interaction. However, no single domain mediated the TRIM71-IMP1 interaction. We performed a CLIP-seq of TRIM71 and discovered many novel RNAs bound by TRIM71 in addition to known RNA targets such as *CDKN1A*. Novel RNA targets included the *MBNL* and *NFI* families as well as developmental and cancer associated genes such as *LIN28B*, *MYC*, *MDM2*, *TCF7*, and *TCF7L2*. Given the interaction between TRIM71 and CRD complex members we looked for overlap of target RNAs finding that 30% of the TRIM71-bound RNAs are also bound by IMP1. This interaction helps mediate the specificity of TRIM71 binding to RNA that was especially apparent when the RNA-binding domain of TRIM71 was deleted. A

TRIM71-ΔNHL mutant was less enriched for target RNAs but there was higher enrichment for overlapping IMP1 targets. We also performed a CLIP-seq on just the RNA-binding domain of TRIM71 finding that it binds to more RNAs and is promiscuous compared to full-length TRIM71. This suggests that protein binding partners do in fact influence RNA binding specificity of TRIM71. Finally, we performed CLIP-seq of TRIM71 NHL point mutants which are associated with congenital hydrocephalus finding that these mutations decrease the enrichment of TRIM71 target RNAs.

TRIM71 binds to CRD complex members

To better understand the function of TRIM71 within a cell we analyzed an unpublished Trim71 immunoprecipitation (IP) mass spectrometry (MS) dataset from Richard Gregory's laboratory performed in mouse embryonic stem cells (mES). Previously, Trim71 was shown to bind to Ago2 in an RNA dependent manner by performing immunoprecipitation in the presence or absence of RNase (Chang et al., 2012). From the IP-MS we sorted bound proteins by peptide count (Figure 2.1). To identify functions of Trim71 and complexes that it might be a member of we ran a gene ontology (GO) analysis of the proteins pulled down in the IP-MS. Gene ontology terms were first analyzed and sorted by False Discovery Rate (FDR) (Figure 2.2A). Based on its binding partners Trim71 was predicted to have a role in biological processes related to RNA processing, translation, etc. Possible molecular functions of Trim71 binding partners included RNA binding and structural components of the ribosome. Cellular component analysis showed that Trim71 was binding to many components within ribonucleoprotein complexes and the cytosolic ribosome. These processes and the binding to many ribosomal components fit with the known functions of Trim71 as an RNA binding protein. Notably, GO terms involved with ubiquitination or related processes were absent in the analysis.

Next, we sorted the GO terms by hierarchy. This accounts for GO terms that are associated with each other and ranks them more highly based on how the groups are enriched. It also accounts for the number of proteins in the dataset associated with each term and the false discovery rate. Hierarchical ranking of categories showed that "CRD Mediated mRNA Stabilization" was the top category in biological processes and cellular components (Figure 2.2B).

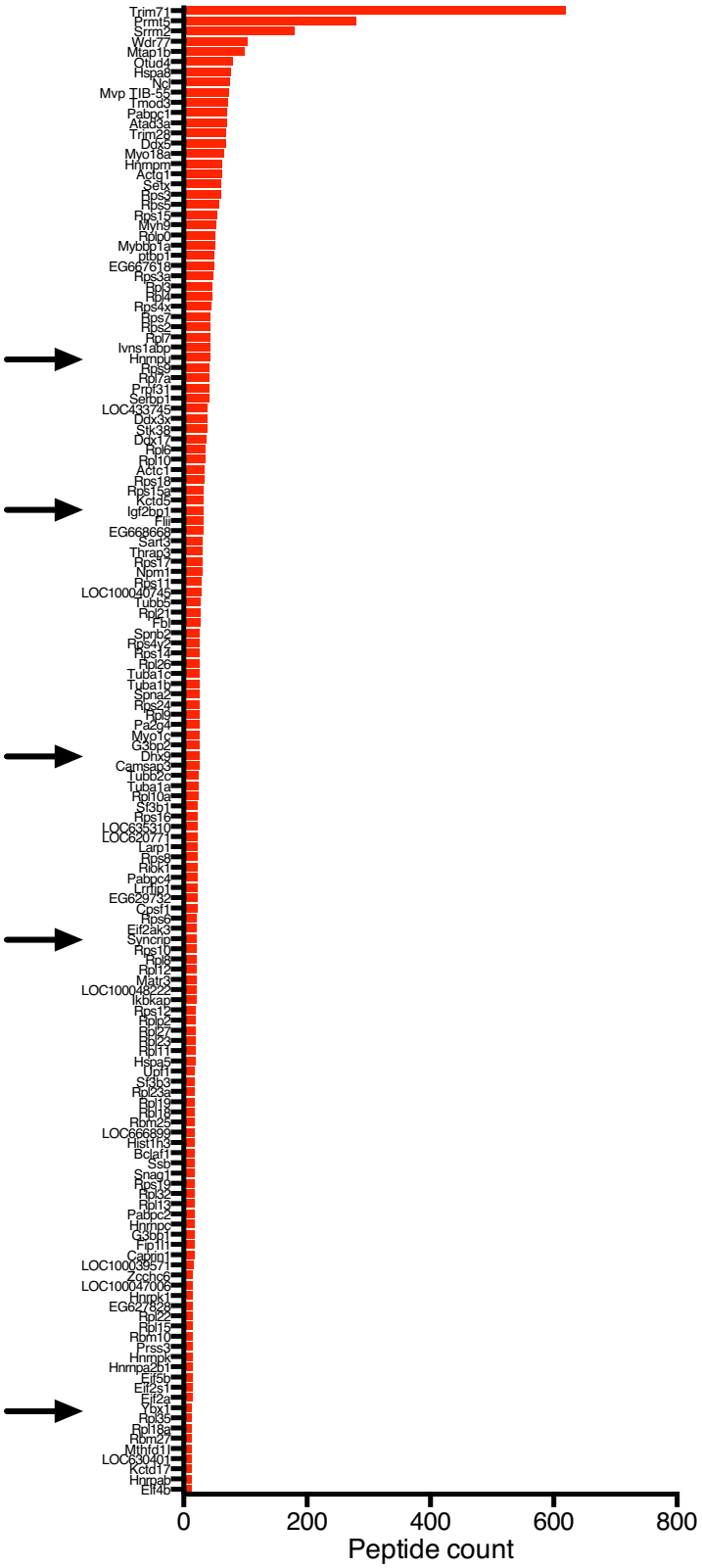
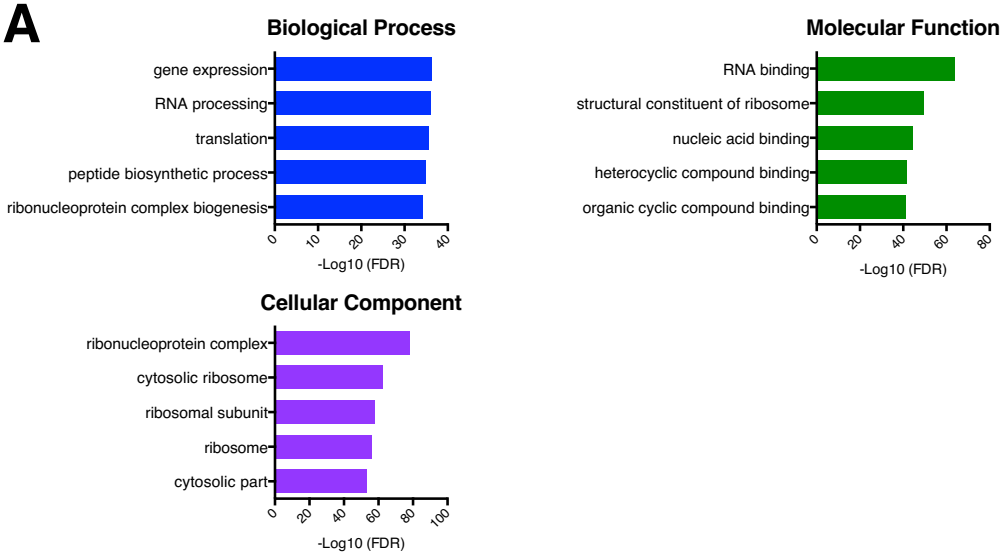


Figure 2.1. TRIM71 IP-MS peptides bound.
All peptides bound in TRIM71 IP-MS in mES cells. The following CRD complex members are noted by arrows: HNRNPU, IGF2BP1 (IMP1), DHX9, SYNCRIP, YBX1.

A



B

Top Biological Processes by Hierarchy

Category	FDR	proteins in dataset/ proteins in category
CRD-mediated mRNA stabilization	2.61E-07	5 / 5
→ mRNA stabilization	3.11E-11	10 / 37
→ RNA stabilization	7.27E-11	10 / 41

Top Molecular Functions by Hierarchy

5.8s RNA binding	6.51E-04	3 / 4
large ribosomal subunit rRNA binding	3.80E-05	4 / 6
regulatory region RNA binding	9.71E-04	3 / 5
ubiquitin ligase inhibitor activity	1.39E-03	3 / 6
mRNA CDS binding	3.67E-02	2 / 5

Top Cellular Components by Hierarchy

CRD-mediated mRNA Stability complex		5 / 6
→ cytoplasmic part	6.78E-17	97 / 8233
Cytosolic small ribosomal subunit	9.35E-30	22 / 55
→ cytosolic ribosome	2.43E-63	48 /152

Figure 2.2. Gene ontology terms of peptides from TRIM71 IP-MS.
(A) Gene ontology terms ranked by significance of all proteins bound by TRIM71. Top hits in all categories included ribosomal components as expected for an RNA-binding protein. **(B)** Gene ontology terms ranked by hierarchy of terms. Top hits in biological process and cellular component include the CRD complex and mRNA stabilization. Top hits by molecular function include ribosomal binding and regulatory region RNA binding.

Furthermore, other mRNA stabilization GO terms were also ranked as significantly enriched. Molecular functions enriched include “Regulatory Region RNA Binding, mRNA CDS Binding, as well as mRNA 5’ UTR and mRNA 3’ UTR Binding.”

The Coding-Region instability Determinant (CRD) is a regulatory element within mRNAs such as *MYC* that the CRD complex binds to in order to promote mRNA stabilization. This complex is composed of IMP1 (IGF2BP1), HNRNPU, SYNCRIP, YBX1, and DHX9 (Weidensdorfer et al., 2009). All 5 proteins were found to bind Trim71 in the IP-MS (Figure 2.1). These proteins, all of which are RNA binding proteins (RNABP), also play roles in other complexes besides the CRD. Additional proteins such as STAU1 and HNRNPE2 are sometimes associated with the CRD complex. Neither of these proteins were present in the Trim71 IP-MS.

Hierarchical ranking also showed enrichment of GO terms associated with ubiquitination processes such as Ubiquitin Ligase Inhibitor Activity and Ubiquitin Protein Ligase Binding (Figure 2.2B). Since TRIM71 is an RNABP it is possible that it is ubiquitinating other RNABPs bound to the same RNA. However, it is still unclear as to what the ubiquitination targets of TRIM71 are.

In addition to the CRD complex, top bound proteins in the IP-MS include Prmt5, an arginine methyltransferase that methylates histones, transcription elongation factors, as well as p53 (Stopa et al., 2015). Another top hit, Wdr77, is an androgen receptor (AR) coactivator known to form a complex with Prmt5. WDR77 is thought to mediate the interaction of the PRMT5 methylosome with its substrates. In addition, as an AR coactivator WDR77 was found play a role in prostate cancer with the protein localizing to the nucleus in normal prostate epithelium but localizing to the cytoplasm in neoplastic and prostate cancer cells (Gao and Wang, 2012). The PRMT5-WDR77 complex methylates

components of the spliceosome as well as histones (Stopa et al., 2015). It is not clear whether TRIM71 could play a role in this complex or whether it regulates these proteins.

As the IP-MS was carried out in mouse ES cells we next wanted to confirm whether interactions observed were conserved in human cells. Using HepG2 cells, we overexpressed an HA-tagged *TRIM71* and performed an IP. We confirmed that TRIM71 was able to co-precipitate IMP1 and DHX9 but not HNRNPU (Figure 2.3A). To supplement this, we then performed an IP of IMP1 and confirmed that IMP1 co-precipitates with TRIM71 and DHX9 but not HNRNPU. Despite the expression of HNRNPU in these cells it seems likely that it is not a key member of the CRD complex due to the lack of co-precipitation with IMP1. Likewise, the absence of STAU1 and HNRNPE2 from the TRIM71 IP suggest that complex members may be absent in various cell types. To test whether these proteins were directly interacting or simply present on the same RNA, half of the samples were treated with an RNase cocktail. RNase treatment did not prevent TRIM71, IMP1, and DHX9 from co-precipitating indicating that their interactions are not mediated by RNA (Figure 2.3A).

TRIM71 domains responsible for interaction with the CRD complex members

Next we tested which domains of TRIM71 were responsible for mediating these protein interactions. Deletion constructs included deletions of each domain on its own, ΔRING, ΔB-box, ΔCoiled-coil, ΔFilamin, and ΔNHL, as well as an NHL-only construct which is deleted for all domains except for the NHL domain. The final construct tested, 7CtoA, had 7 point mutations within the RING domain which ablated ubiquitination enzymatic activity but otherwise left the domain intact. All TRIM71 constructs were

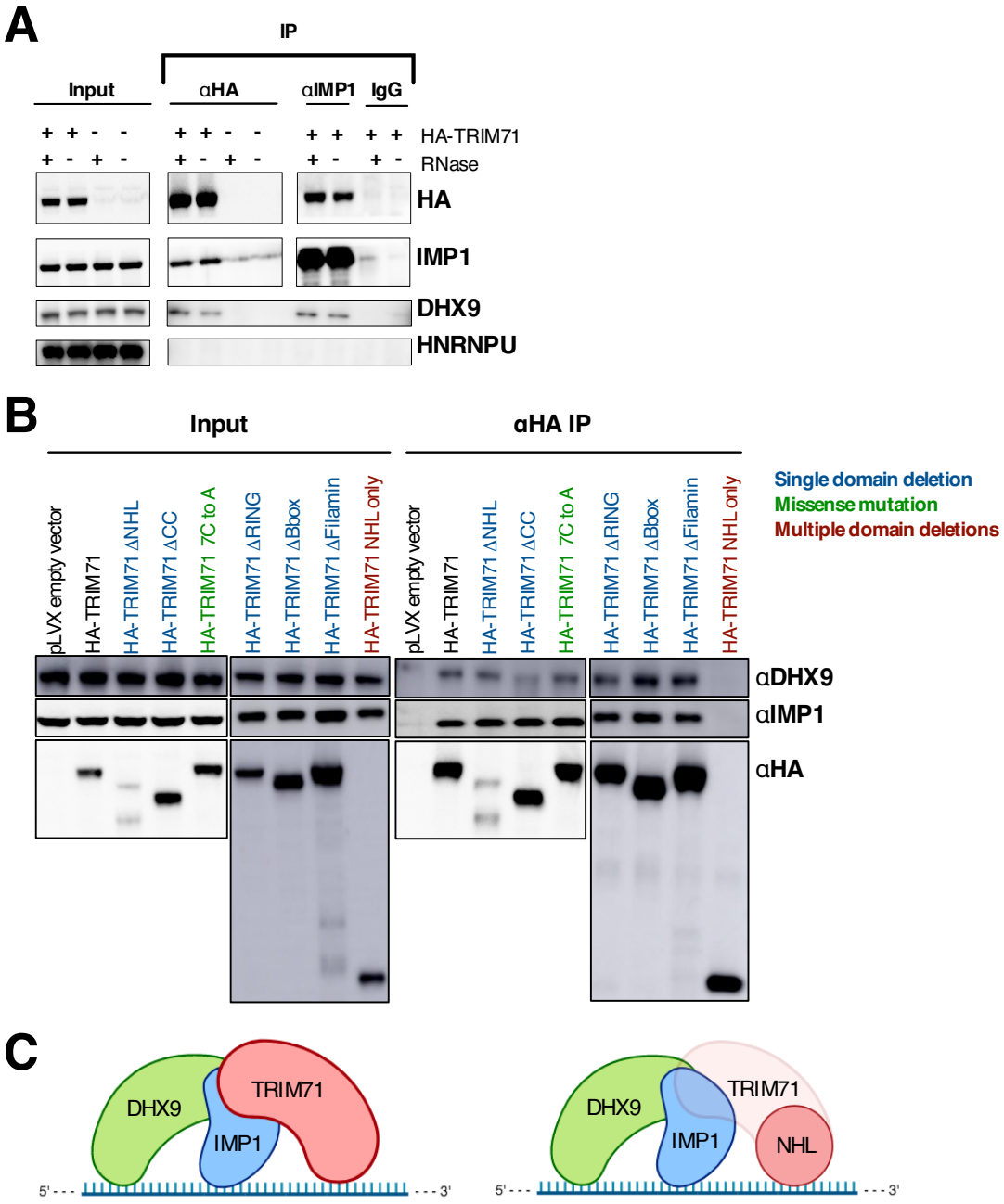


Figure 2.3. TRIM71 co-IPs with CRD complex members IMP1 and DHX9.
(A) TRIM71 co-IPed with CRD complex members IMP1 and DHX9. IMP1 also co-IPed with TRIM71 and DHX9. Neither TRIM71 nor IMP1 co-IP with HNRNPU. The presence of RNase did not impact the ability of any proteins to co-IP each other. **(B)** TRIM71 domain deletion analysis to test for domains necessary to co-IP IMP1 and DHX9. No single domain was necessary to bind IMP1 while the coiled-coil domain was necessary to bind DHX9. The NHL-only mutant that lacks all protein binding domains failed to co-IP IMP1 and DHX9. **(C)** Model of how TRIM71 binds to CRD complex members, left. The NHL domain on its own is unable to bind the CRD complex, right. The translucent portions of TRIM71 represent the domains deleted in part B.

precipitated in the presence of RNase in order to test specifically for protein-protein interactions not mediated by RNA. First testing for co-IP with DHX9, the ΔCoiled-coil and NHL-only constructs show weaker interaction. The coiled-coil domain is a protein binding domain and it appears to mediate the interaction with DHX9.

Next testing for co-IP with IMP1, each construct except for NHL-only was able to co-IP IMP1 (Figure 2.3B). This would indicate that IMP1 and TRIM71 may interact along several protein domains and deletion of any one domain does not prevent their interaction. Alternatively, the interaction could still be mediated by RNA however they would have to be binding the RNA very close together such that RNase is unable to access it. The last possibility is that there is an unannotated protein interaction domain of TRIM71. There is a 100 amino acid region between the RING and B-box domains that does not have an annotated domain or function. However, this long stretch of amino acids, rich in histidine and proline, could account for the interaction. We did not test this region for functionality or binding to IMP1 but it remains a possibility.

The NHL domain on its own does not interact with either IMP1 or DHX9. As this domain is the known RNA binding domain of TRIM71 it makes sense that it is not responsible for protein-protein interactions. A model for TRIM71 binding to the CRD complex on RNA is shown in Figure 2.3C. Full-length TRIM71 is able to bind both members of the CRD complex whereas the NHL domain is only binding RNA and does not interact with CRD complex members.

TRIM71 crosslinking IP sequencing shows binding to a large number of RNAs

The majority of TRIM71 associated phenotypes from *C. elegans* to humans arise from mutations in the RNA-binding NHL domain (Furey et al., 2018; Slack et al., 1999).

This suggests that the RNA binding activities are crucial to its functions in a cell. However, the RNAs bound by TRIM71 have not been studied on a global scale. We profiled these RNAs via crosslinking IP (CLIP) because crosslinking of the RNA to protein allows for harsh IP wash conditions that are necessary to wash away highly expressed and non-specifically bound RNAs. Non-crosslinked RNA IPs often have high levels of non-specifically bound RNAs. To aid in RNA-crosslinking we incubated cells in the presence of 4-thiouridine (4-SU) an ultraviolet-sensitive uridine analog that incorporates into newly synthesized RNA. When cells are treated with UV light at 365nM, the photoactivatable ribonucleotide (PAR) irreversibly crosslinks to protein. It is possible to crosslink RNA without 4-SU at higher energies with 254nM UV light, however the efficiency of crosslinking is low. We initially attempted CLIP with higher energy UV without PAR but the signal was undetectable in CLIP-qPCR experiments of known RNA targets (data not shown).

Although CLIP experiments involve many harsh wash steps, the technique is not perfect in determining RNAs bound by particular proteins. There are challenges that include non-specific binding of highly expressed RNAs. However, there are ways that this can be partially be addressed. For example, with PAR-CLIP there are predictable thymine to cytosine or adenine to guanosine transitions that occur where a protein was crosslinked to a 4-SU nucleotide. These transitions can be looked at as a ratio of other transitions that occur to determine where a protein could be binding. Other types of CLIP experiments, besides PAR-CLIP, rely on transitions to predict the binding site but they do not cause the predictable T to C or A to G transitions like the photoactivatable ribonucleotide, making it difficult to sort out the true signal from the noise.

PAR-CLIP experiments were carried out in HepG2 cells due to the strong phenotypes associated with *TRIM71* knockdown in these cells (data shown in Chapter 3). In most CLIP studies, the immunoprecipitated RNA is sequenced and top RNA hits are determined based on which RNAs are bound most strongly. However, this analysis relies on the RNA being highly expressed. Therefore, we added input samples, taken before the IP, to our sequencing analysis to account for expression differences of the RNAs.

Each input and IP sample was sequenced in duplicate and we obtained ~20 million reads per sample. Both replicates of the IP and input were well correlated for all samples with a Pearson coefficient $r>0.95$ (Figure 2.4). From here we performed a differential expression analysis using DEseq to determine what RNAs were enriched in the IP compared to the input. There were 2432 RNAs significantly enriched by greater than a fold change of 2 in the IP samples (Figure 2.5A). To check whether highly expressed RNAs were more likely to be enriched in the IP we plotted the base mean expression of all 4 samples versus the fold change of the TRIM71 IP (Figure 2.5B). The distribution of enrichment versus expression was not shifted to suggest that expression of RNAs correlated with IP enrichment. This indicates that the CLIP washes were likely successful in washing away non-specific, highly expressed RNAs. Lowly expressed RNAs, below a base mean expression of around 15 counts, were easily skewed and showed large differences in expression between samples. Thus, as is typical for differential expression analysis, very lowly expressed RNAs were excluded from further analysis.

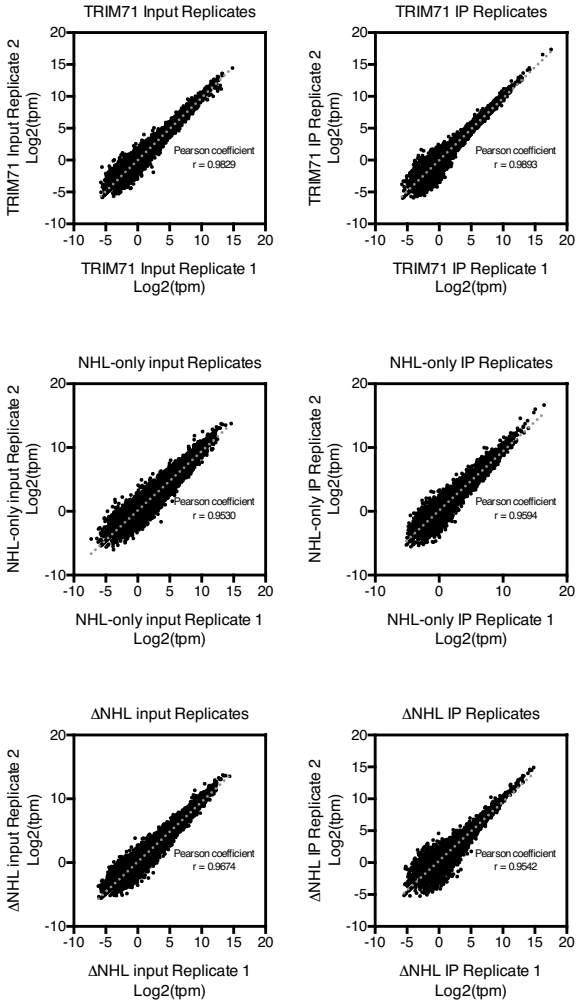


Figure 2.4. CLIP sequencing replicates
Replicates 1 and 2 were well correlated for all samples with Pearson coefficients greater than 0.95.

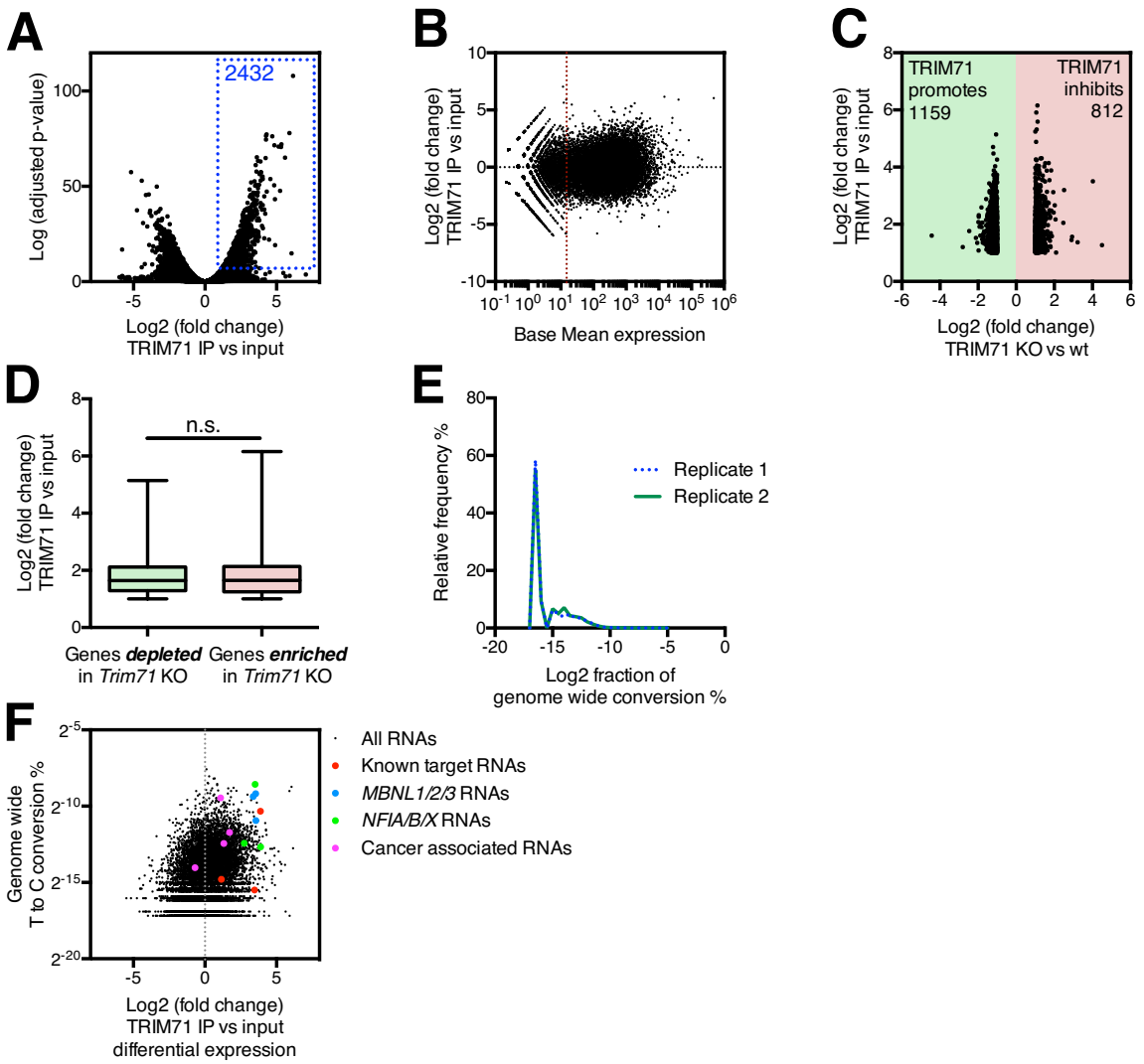


Figure 2.5. TRIM71 CLIP-sequencing.

(A) Differential expression analysis of IP vs Input samples. There were 2432 RNAs significantly enriched greater than a log2 fold change of 1 in the IP samples compared to the input (shown in the blue box). These also have expression values greater than a base mean of 15. **(B)** Base mean expression vs fold change confirms that highly expressed genes are not overrepresented in the IP samples. Only RNAs with a base mean over 15 (red line) were called as hits. **(C)** Of the 2432 RNAs enriched in the IP, 1971 of these were present in a published RNA-seq data from Mitschka et al., 2015 of *Trim71* KO mouse ES cells. 1159 TRIM71-bound RNAs were decreased by greater than a log2 fold change of 1 in the *Trim71* KO cell lines indicated that Trim71 was promoting these RNAs. 812 RNAs were increased, indicating that Trim71 was inhibiting these. **(D)** Genes enriched or depleted in the *Trim71* KO mES cells were not bound differently in the TRIM71 IP. Mann-Whitney test, p=0.86. **(E)** Frequency distribution of T to C conversions throughout the genome. Most genes have no conversions while a small % have many conversions. **(F)** T to C conversion % per gene is higher in IP enriched genes. Colored dots represent shown genes followed up on with CLIP-qPCR.

We next wanted to determine how TRIM71 regulated RNAs on a global level. Using a published RNA-seq dataset comparing *Trim71* knockout versus wild-type mouse ES cells, we looked at whether directly bound RNAs from the CLIP experiment were enriched or depleted in *Trim71* KO cells (Figure 2.5C). Only around 80% of the RNAs in our dataset were profiled in the mES RNA-seq. From these, around 60% of the RNAs were depleted in *Trim71* KO cells which indicates that TRIM71 could be positively regulating these RNAs. The other 40% were enriched in the *Trim71* KO cells which could indicate negative regulation. This result is slightly surprising given that the only described role of TRIM71 as an RNABP is to negatively regulate RNAs. However, as I have already shown, TRIM71 is able to bind CRD complex members that are known to positively regulate RNAs. Since these protein interactions are conserved from at least mice to humans, we find that these results support our hypothesis that TRIM71 can positively regulate RNAs, likely as a member of the CRD complex. Our results still support past studies showing that TRIM71 is able to negatively regulate RNAs, however TRIM71 appears to be capable of bidirectional regulation. Next, we tested if the genes inhibited or promoted were bound more strongly by TRIM71 in the CLIP to understand whether the strength of TRIM71 binding influences its effect on an RNA. Although we did not directly test binding strength or compare binding affinities between RNAs, we were able to look at enrichment of RNA in the TRIM71 IP. Looking at enrichment in the IP based on positive or negative regulation of the target, we did not see a significant difference (Figure 2.5D). This suggests that TRIM71 does not favor inhibition or promotion of target RNAs, at least in terms of enrichment in the CLIP.

Given that PAR-CLIP causes thymine to cytosine transitions, we next analyzed these transitions in our dataset. We looked at the frequency distribution of T to C

throughout the genome by calculating the number of transitions for every gene then determining what this represented as a fraction of all T to C transitions in the dataset (Figure 2.5E), accounted for differences in gene length. We plotted the fraction of genome wide conversion percentages as a frequency distribution and showed that most genes had very few (zero or one) transitions, as represented by the large leftward peak. There were a small number of genes that represented a larger percentage of transitions in the data set as shown by the smaller peaks towards the middle of the plot.

Finally, we plotted the fraction of genome wide conversion percentage for each gene against the fold change enrichment from the differential expression analysis (Figure 2.5F). The rightward shift in the distribution shows that genes more strongly enriched in the IP represent a larger percentage of the T to C transitions in the dataset. This indicates that the CLIP protocol sufficiently enriched for bound RNAs. The colored points are specific genes we chose to follow up on with CLIP-qPCR and functional studies.

Gene Ontology analysis of RNAs from the TRIM71 CLIP

In order to uncover potential biological roles of TRIM71 regulation on RNA we performed a gene ontology analysis of the RNAs pulled down in the CLIP. Knowing which RNAs were likely to be inhibited or promoted was important for this analysis as it influences how we think about the GO terms with regards to regulation by TRIM71. Therefore, we separated the bound RNAs into groups as previously described based on whether the genes were likely to be inhibited or promoted. The 1159 RNAs thought to be promoted by TRIM71 were analyzed separately from the 812 that may be inhibited by TRIM71.

GO analysis suggests that biological processes possibly inhibited by TRIM71 include “macrophage differentiation, DNA damage response, and signaling involved in mitotic checkpoints” (Figure 2.6). Molecular processes possibly inhibited include “cyclin dependent kinase activity.” Some of these terms fit within the literature of what is known about TRIM71 and its involvement in early development and induced pluripotency. Other terms such as “spliceosomal assembly” would indicate that TRIM71 is possibly regulating novel RNAs not previously discovered.

Biological processes possibly promoted by TRIM71 include “aspartate transport, histone methylation, and blastocyst growth.” Molecular processes include “histone methyltransferase and deubiquitinase activity.” Like above, some of these terms fit well within the literature such as “blastocyst growth” given the role of TRIM71 in early development. Other terms such as “histone methylation” do not fit in the context of what is known about TRIM71.

TRIM71 binds to known and novel RNAs

Published TRIM71 targets in cells where TRIM71 is biologically important include *CDKN1A* and *EGR1* (Chang et al., 2012; Rand et al., 2018; Worringer et al., 2014). These targets were discovered in ES and iPS cells and were experimentally validated as targets using RNA IP followed by qPCR and luciferase reporter assays for *CDKN1A*. The regulation of these RNAs by TRIM71 was further validated at the protein level. TRIM71 has also been studied in HEK293 cells where it is not endogenously expressed

A

Biological processes inhibited by TRIM71	Fold Enriched	FDR
positive regulation of macrophage differentiation (GO:0045651)	8.7	3.73E-02
mitochondrial translational elongation (GO:0070125)	6.23	3.16E-07
spliceosomal snRNP assembly (GO:0000387)	6.18	5.02E-03
translational termination (GO:0006415)	5.65	1.03E-06
mitochondrial translational termination (GO:0070126)	5.57	5.52E-06
DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest (GO:0006977)	5.4	1.35E-03
mitochondrial translation (GO:0032543)	5.36	5.81E-07
signal transduction involved in mitotic G1 DNA damage checkpoint (GO:0072431)	5.31	1.52E-03
signal transduction involved in mitotic cell cycle checkpoint (GO:0072413)	5.31	1.51E-03
signal transduction involved in mitotic DNA integrity checkpoint (GO:1902403)	5.31	1.49E-03

Molecular processes inhibited by TRIM71	Fold Enriched	FDR
cyclin-dependent protein serine/threonine kinase activity (GO:0004693)	7.2	8.83E-03
cyclin-dependent protein kinase activity (GO:0097472)	6.96	1.04E-02
structural constituent of ribosome (GO:0003735)	4.4	3.17E-07
DNA-binding transcription repressor activity, RNA polymerase II-specific (GO:0001227)	3.14	1.52E-04
proximal promoter sequence-specific DNA binding (GO:0000987)	2.16	8.91E-04
RNA polymerase II proximal promoter sequence-specific DNA binding (GO:0000978)	2.12	2.72E-03
DNA-binding transcription activator activity, RNA polymerase II-specific (GO:0001228)	2.05	3.15E-02
DNA-binding transcription activator activity (GO:0001216)	2.05	3.05E-02
RNA polymerase II regulatory region sequence-specific DNA binding (GO:0000977)	2.05	1.51E-04
RNA polymerase II regulatory region DNA binding (GO:0001012)	2.03	1.64E-04

B

Biological processes promoted by TRIM71	Fold Enriched	FDR
aspartate transmembrane transport (GO:0015810)	10.04	2.94E-02
histone mRNA catabolic process (GO:0071044)	9.03	1.49E-02
histone H3-K9 methylation (GO:0051567)	7.23	3.09E-02
blastocyst growth (GO:0001832)	6.66	1.86E-02
histone H3-K9 modification (GO:0061647)	6.66	1.85E-02
regulation of centriole replication (GO:0046599)	5.75	3.27E-02
protein K63-linked deubiquitination (GO:0070536)	5.25	1.27E-02
protein K48-linked deubiquitination (GO:0071108)	5.08	1.48E-02
tRNA-containing ribonucleoprotein complex export from nucleus (GO:0071431)	4.78	1.99E-02
tRNA export from nucleus (GO:0006409)	4.78	1.98E-02

Molecular processes promoted by TRIM71	Fold Enriched	FDR
histone methyltransferase activity (H3-K9 specific) (GO:0046974)	9.03	3.97E-02
Lys63-specific deubiquitinase activity (GO:0061578)	9.03	1.58E-02
translation repressor activity (GO:0030371)	5.56	1.96E-02
protein serine/threonine/tyrosine kinase activity (GO:0004712)	4.2	2.39E-02
ubiquitin-like protein-specific protease activity (GO:0019783)	3.59	7.21E-05
ribosome binding (GO:0043022)	3.55	2.40E-02
histone acetyltransferase activity (GO:0004402)	3.39	3.13E-02
peptide N-acetyltransferase activity (GO:0034212)	3.37	1.56E-02
peptide-lysine-N-acetyltransferase activity (GO:0061733)	3.29	3.80E-02
ubiquitinyl hydrolase activity (GO:0101005)	3.27	1.28E-03

Figure 2.6. Gene Ontology analysis of RNAs TRIM71 promotes or inhibits.
(A) Top biological and molecular processes possibly inhibited by TRIM71 included differentiation, splicosomal assembly, cell cycle checkpoints, and specific transcription factors. **(B)** Top processes possibly promoted by TRIM71 included blastocyst growth, histone methylation and modifications, and deubiquitination.

(Loedige et al., 2013). The main TRIM71 target found and validated in HEK293 cells was *CCNE2* and *E2F7* (Loedige et al., 2013). Unfortunately, neither of these targets were validated at the protein level so it is unclear exactly what effect TRIM71 has. In summary, the known RNA targets of TRIM71 seem to be experimentally and cell type dependent. The target with the strongest evidence to date is *CDKN1A* as it has been repeated several times.

As previously mentioned, our CLIP experiments were carried out in HepG2 cells which have strong phenotypes associated with *TRIM71* knockdown (data shown in Chapter 3). We thought it was important to carry out the CLIP in cells showing a phenotype to assure that biologically relevant target RNAs were discovered and to enable validation through genetic experiments. In cells without a TRIM71 phenotype, it is possible that less biologically relevant RNAs would be bound; and as shown above with HEK293 cells, validation experiments prove difficult. Furthermore, it is also possible that protein binding partners required for the action of TRIM71 would be absent in cells without a phenotype leading us to question the relevance of bound RNAs.

In the TRIM71 CLIP performed in HepG2 cells we confirmed that *CDKN1A*, *EGR1*, *CCNE2*, *HOXA5*, *ZNF362*, and *E2F7* were targets of TRIM71 significantly enriched in the IP versus input samples. Gene tracks of *CDKN1A* and *EGR1* showed enrichment of reads in the IP samples compared to the input (Figure 2.7A) with *CDKN1A* being one of the most enriched RNAs in the IP versus input differential expression analysis. Additionally, *CDKN1A* had a large number of T to C transitions whereas *EGR1* and *CCNE2* did not, likely because they were quite lowly expressed in the samples (Figure 2.5F, red points). These results fit with previous low-throughput studies looking at TRIM71 targets suggesting that the CLIP sufficiently enriched for RNAs of true targets.

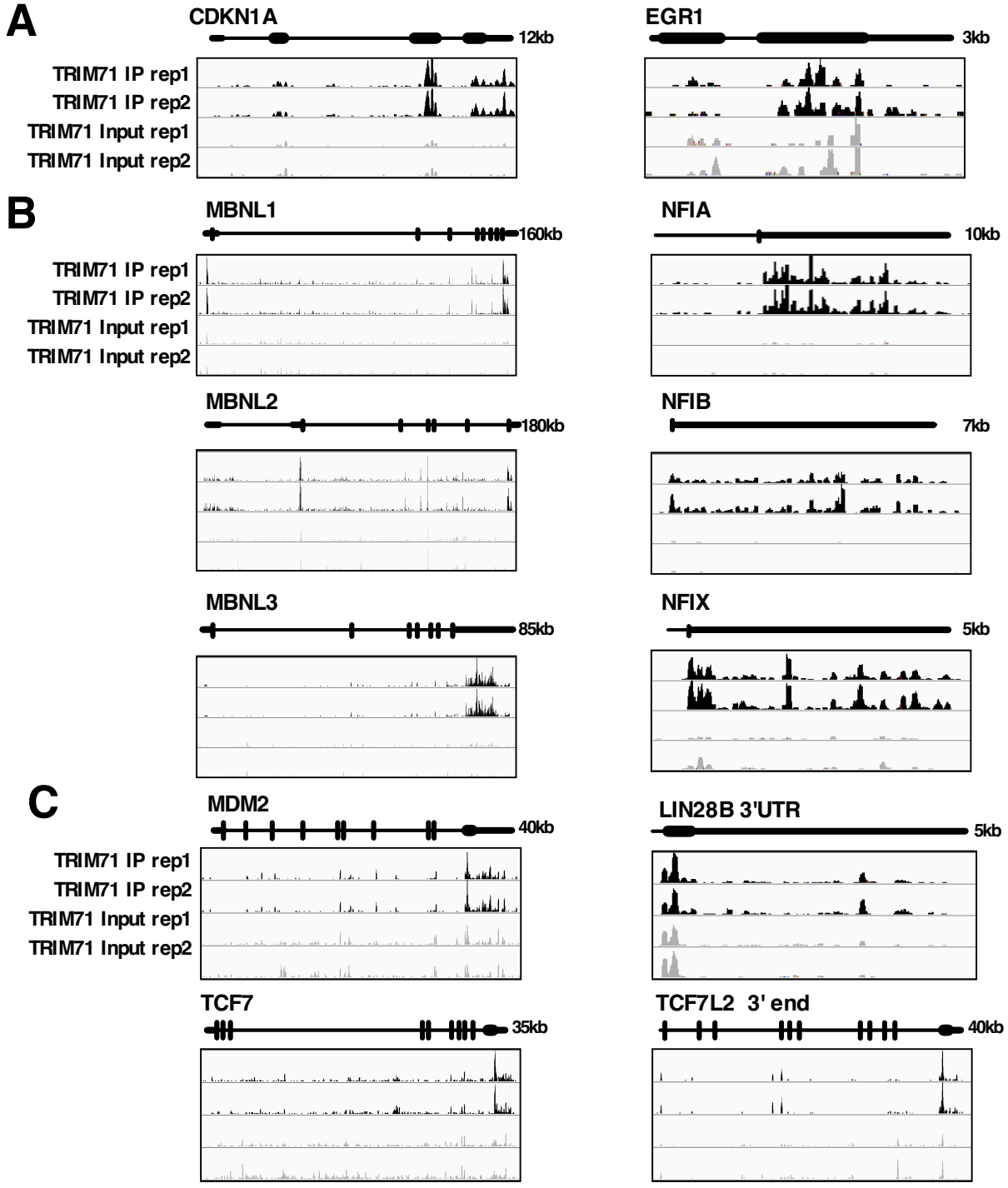


Figure 2.7. Gene tracks of known and novel TRIM71 targets.
(A) Known TRIM71 target mRNAs, *CDKN1A* and *EGR1* were both significantly enriched in the TRIM71 CLIP. **(B)** All three members of the MBNL family (1/2/3) and three of four members of the NFI family (A/B/X) were significantly enriched in the TRIM71 CLIP and were among the most highly enriched. **(C)** Additional interesting target mRNAs enriched in the CLIP included *MDM2*, *LIN28B*, *TCF7*, and *TCF7L2*. All genes are shown in left to right orientation.

From here we went about analyzing the TRIM71 targets in several ways in order to prioritize what to follow up on. First, we looked at whether there were any gene families where TRIM71 was binding several members. We found that the *MBNL* and *NFI* families both had most members bound by TRIM71. Second, we looked at cancer associated genes and found targets such as *LIN28B*, *MDM2*, *IDH1*, *BRCA1* and *BRCA2*, *TCF7* and *TCF7L2*, *E2F6*, *EGFR*, and *ATM*. Third, we looked for overlap of TRIM71 bound RNAs with IMP1 bound RNAs and found many genes already mentioned in addition to *MYC*.

The MBNL family

The muscleblind-like (*MBNL*) family is composed of 3 members, *MBNL1/2/3*. These are RNA-binding proteins that act as tissue specific splicing regulators. *MBNL1/2* expression increases during development and induces adult specific splicing patterns on a global scale (Konieczny and Stepniak-Konieczna, 2014). *MBNL1* is expressed in a wide variety of adult tissues however mouse phenotypes suggest its predominant effect is in adult muscle tissue (Sznajder et al., 2016). Likewise, *MBNL2* is also expressed in a wide variety of adult tissues however expression is highest in the brain and mouse phenotypes suggest an importance in neural development (Charizanis et al., 2012; Sznajder et al., 2016). Meanwhile *MBNL3* has tissue expression patterns more localized to testis, lung, spleen, and placenta. In the liver, *MBNL3* is expressed fetally and is turned off during development. However it can become re-expressed in hepatocellular carcinoma (Yuan et al., 2017). The MBNL family is most closely associated with myotonic dystrophy, a genetic disease of degenerative muscle weakness. The mechanism of this disease involves

nucleotide repeats within the *DMPK* mRNA that sequesters/sponges MBNL proteins preventing their regulatory effects on splicing (Lee and Cooper, 2009).

The TRIM71 CLIP-seq showed binding and enrichment of each *MBNL* family member. All three were among the top 100 most enriched RNAs. Plotting the gene tracks show binding within the 3'UTRs as well as the coding sequences of these mRNAs (Figure 2.7B). The strong enrichment of reads within the IP but absent in the input is also apparent. Furthermore, all three family members were highly enriched for T to C transitions (Figure 2.5F, blue points). Gene ontology analysis suggested that "spliceosomal snRNA assembly" was one of the top processes inhibited by TRIM71, and these three family members, as well as other mRNAs from the CLIP, were contained within this GO term (Figure 2.6A).

Based on published data from mES *Trim71* knockout cells, TRIM71 appears to inhibit all three mRNAs as they were enriched in the *Trim71* KO cells (Mitschka et al., 2015). Since *Trim71* knockout mice and patients with *TRIM71* mutation exhibit severe neural defects (Furey et al., 2018; Schulman et al., 2008), it is possible that this could be due to loss of regulation of the *MBNL* family, or more specifically *MBNL2*. The hypothesis would be that early in development TRIM71 inhibits *MBNL2* and thereby inhibits adult splicing patterns. Loss or mutation of *TRIM71* could lead to premature expression of *MBNL2* and premature differentiation of neurons which could explain the mouse and human brain phenotypes where cells appear to prematurely differentiate (Chen et al., 2012; Mitschka et al., 2015).

The *MBNL* family is conserved from *C. elegans* to human, with worms having one member, *mbl-1*. The functions are also relatively conserved, with *mbl-1* loss resulting in muscle defects as well as synapse defects of the neuromuscular junction (Spilker and

Wang, 2012). Though *lin-41* has not been connected to *mbl-1* in *C. elegans* it is reasonable to hypothesize that they may be connected. *lin-41* loss leads to misexpression of muscle myosins and also has an effect on neuronal regeneration (Tocchini et al., 2014; Zou et al., 2013).

The NFI family

The nuclear factor 1 (*NFI*) family is composed of 4 members, *NFIA/B/C/X*. These are transcription factors that dimerize and are capable of enhancing or repressing gene expression. These transcription factors regulate a vast array of genes from those that are ubiquitously expressed to those involved in hormonal signaling and development (Gronostajski, 2000). Each member is subject to alternative splicing, having many possible isoforms, and in fact *NFIX* is a major target of MBNL splicing regulation (Konieczny and Stepniak-Konieczna, 2014). *NFIA/B/X* are all highly expressed in the developing brain, specifically within the ventricular zone and cortical plate (Plachez et al., 2008), areas of active neuronal cell division and migration during brain development. These three family members have severe effects when deleted in mice. *Nfia/b* mutants have enlarged ventricles and other brain defects with *Nfia* deletion also causing hydrocephalus (Heng et al., 2012). As a result of these and other developmental defects, *Nfia* and *Nfib* mutant mice do not survive birth. *Nfix* has overlapping expression patterns with *Nfia/b* and deletion causes abnormal brain development but it lacks the same lethal phenotype (Heng et al., 2012). *NFIC* is the only family member not involved in brain development. Rather, it plays a role in tooth development and liver regeneration (Edelmann et al., 2015; Kim et al., 2015).

The *NFI* family regulates neuronal development and migration by controlling expression of many neuronal and glial genes such as *GFAP, GLAST,* and cell adhesion molecules (Barry et al., 2008; Heng et al., 2012). The severe mutant phenotypes are mostly due to the failure of neurons to properly differentiate and migrate. Interestingly, although NFI genes promote differentiation during development, in the adult brain NFIX promotes maintenance of neural stem cells (Martynoga et al., 2013).

The TRIM71 CLIP-seq showed binding and enrichment of *NFIA/B/X*. All three were within the top 100 bound genes. TRIM71 binds to the 3'UTRs of these genes (Figure 2.7B) and also within the coding region (not shown). These three genes also had many thymine to cytosine transitions indicating binding within these regions (Figure 2.5F, green points). Gene ontology terms associated with this family includes "DNA-binding transcription repressor activity" and "DNA-binding transcription activator activity", both of which came up as significant in the GO molecular processes analysis (Figure 2.6A).

TRIM71 is expected to inhibit these genes based on their enrichment in *Trim71* knockout mouse ES cells (Mitschka et al., 2015). Furthermore, loss of *Trim71* in mouse ES cells causes an increase in the neuronal development pattern. Due to the strong connections and similar phenotypes in brain development, connecting TRIM71 with the *NFI* family is a reasonable hypothesis. It is possible that TRIM71 inhibits the *NFI* family during early brain development. Loss of *TRIM71* is shown to lead to premature neural differentiation which could be due to early expression of the *NFI* family. Mutation of *TRIM71* leads to hydrocephalus in patients (Furey et al., 2018), possibly due to early differentiation within the brain. Although loss of NFI prevents proper neural differentiation, it ends up leading to similar phenotypes as *TRIM71* loss. If the *NFI* family is a target of

TRIM71 we'd expect precocious *NFI* overexpression to mimic the *TRIM71* mutant phenotype however this has not been tested in mice.

The NFI family is also conserved in *C. elegans* where there is one member, *nfi-1*, that shares some properties with the vertebrate homologs (Lazakovitch and Kalb, 2005). The DNA binding properties of *C. elegans* NFI-1 are most similar to human NFIC. Worms with mutant *nfi-1* exhibit locomotor, egg-laying, and pharyngeal pumping defects which could be related to defects in muscles, neurons, or both (Lazakovitch and Kalb, 2005).

Cancer associated genes

The second method we used to analyze hits from the TRIM71 CLIP-seq was to look for the presence of interesting genes in a targeted manner. Specifically, we looked for cancer associated genes. Some of the cancer associated genes bound by TRIM71 include *LIN28B*, *MDM2*, *IDH1*, *BRCA1* and *BRCA2*, *TCF7* and TCF7L2, *E2F6*, *EGFR*, and *ATM*. All of these genes showed enrichment in the IP versus input differential expression analysis and they all met the expression value cutoffs to be included in the analysis. However, *BRCA1/2*, *EGFR*, and *ATM* were still fairly lowly expressed so we did not follow up on them. Typical for PAR-CLIP analysis, we next looked for the presence of thymine to cytosine or adenine to guanosine transitions in the sequencing reads indicating binding. Although all the remaining genes had these specific transitions, *MDM2*, *LIN28B*, and *TCF7L2* stood out as having the most (Figure 2.5F, magenta points). All three genes were highly enriched in IP sample reads throughout the gene (Figure 2.7C). Based on the depletion of *Lin28b* in mES *Trim71* knockout cells, TRIM71 would be expected to promote

LIN28B. *Mdm2* and *Tcf7l2* levels were not significantly changed in the mES cells (Mitschka et al., 2015).

As described in the introduction, LIN28B and TRIM71 are members of the same genetic pathway important during early development and they correlate in expression. In a similar manner to the interaction between LIN28B and IMP1, TRIM71 could feed back on *LIN28B* to positively regulate its expression. LIN28B has been shown to regulate *IMP1* directly by binding the mRNA and indirectly by inhibiting *let-7* (Nguyen et al., 2014). IMP1, itself an RNA binding protein, feeds back on *LIN28B* and positively regulates it (Conway et al., 2016; Müller et al., 2018). We hypothesized that the same genetic interaction occurs between LIN28B and TRIM71.

MDM2 is highly expressed in HepG2 cells and was strongly enriched in the TRIM71 IP (Figure 2.7C). It is a key regulator of p53 that is capable of ubiquitinating it leading to degradation. As an inhibitor of p53, *MDM2* is an oncogene in many cancers (Wade et al., 2013). Perhaps not surprisingly, *MDM2* is more frequently amplified in p53 wild-type tumors. Besides regulating p53, MDM2 has other p53-independent roles within a cell. They are still pro-growth/oncogenic and include ubiquitination of alternative targets, transcription factor regulation, and mRNA binding (Bohlman and Manfredi, 2014). These mechanisms have effects on the cell cycle, apoptosis, genome stability, and other cell processes (Bohlman and Manfredi, 2014). As MDM2 has been shown to have a functional impact in HepG2 cells (Wang et al., 2011), we thought it was an intriguing target of TRIM71 to follow up on. Based on TRIM71 phenotypes in HepG2 cells (shown in chapter 3) we hypothesized that TRIM71 would promote *MDM2*.

TCF7 and TCF7L2 are transcription factors that play a role in Wnt signaling. When a Wnt ligand binds to a receptor, β-catenin accumulates in the cytoplasm and translocates to the nucleus where it binds to TCF/LEF transcription factors as a co-activator. There are four *TCF/LEF* family members in humans, two of which were bound in the TRIM71 CLIP. *TCF7* was somewhat lowly expressed in HepG2 cells, though it was still enriched in the IP (Figure 2.7C). Due to its low expression it was difficult to evaluate for the presence of thymine to cytosine transitions. *TCF7L2* was highly expressed, highly enriched, and contained many crosslink-induced mutation sites (Figure 2.7C). The other *TCF/LEF* family members, *TCF7L1* and *LEF1*, did not have detectable mRNA expression in HepG2. Since HepG2 is a β-catenin mutant driven cell line, the presence of TCF/LEF transcription factors may be vital for its growth. Chemical disruption of β-catenin/TCF complexes has a dose dependent cytotoxicity in HepG2 (Wei et al., 2010) leading us to follow up on *TCF7/TCF7L2* as potential TRIM71 targets.

Overlap of TRIM71 and IMP1 bound mRNAs

Since I have shown that TRIM71 is capable of binding CRD complex members IMP1 and DHX9, we wanted to test whether there was overlap between the RNAs bound. IMP1 is a well profiled RNA binding protein for which many CLIP-seq experiments have been carried out (Conway et al., 2016; Hafner et al., 2010; Müller et al., 2018; Van Nostrand et al., 2016). IMP1 CLIP in ES cells showed that 3304 RNAs were significantly enriched in the IP compared to input based on a differential expression analysis (Conway et al., 2016). Comparing this to the 2432 RNA enriched in the TRIM71 CLIP, we found 735 RNAs were bound by both proteins (Figure 2.8A). This represents around 30% of the

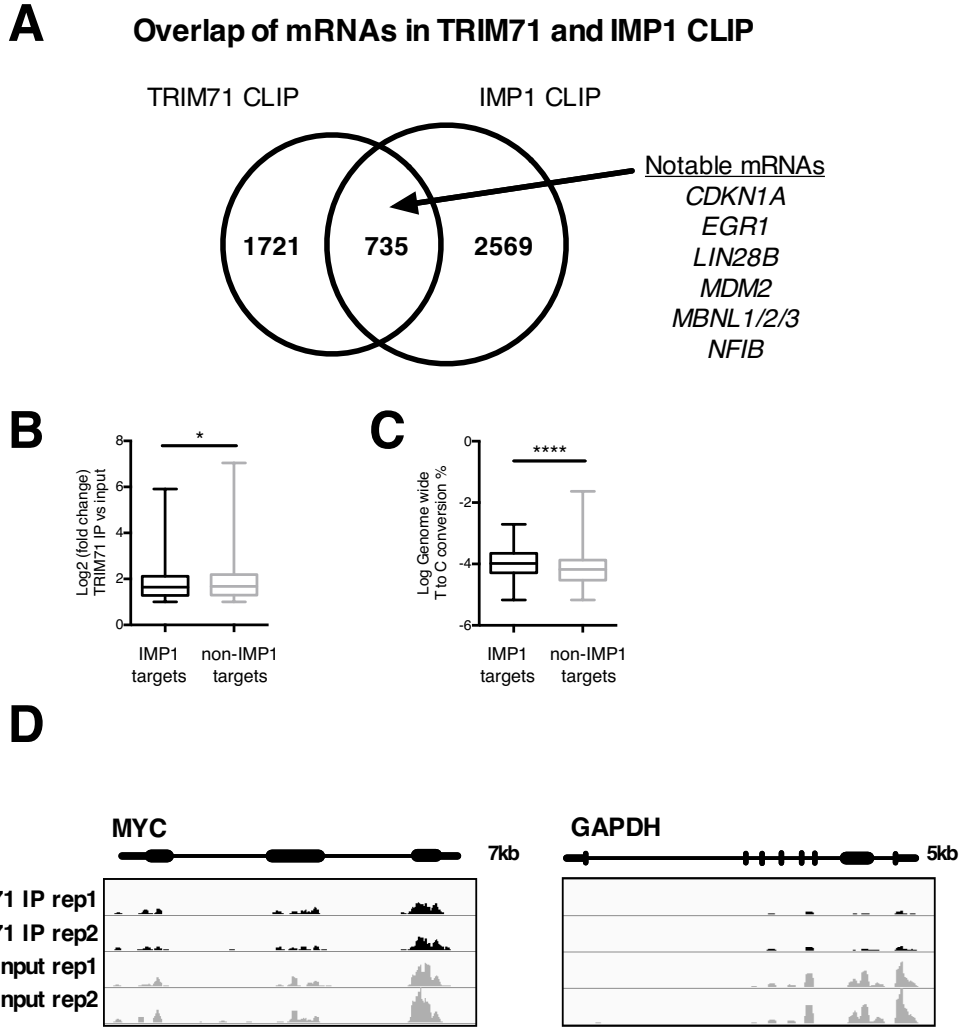


Figure 2.8. Overlap of TRIM71 and IMP1 bound RNAs
(A) Of the 2456 TRIM71-bound RNAs, 30% of them were also bound by IMP1 based on the published IMP1 CLIP dataset from Conway et al., 2016. **(B)** TRIM71 target RNAs also bound by IMP1 were slightly less enriched than non-IMP1 target RNAs. *t*-test p=0.037.**(C)** IMP1 target RNAs represented a larger % of T to C transition than non-IMP1 targets. *t*-test p<0.0001. **(D)** Canonical IMP1 target, *MYC*, was not significantly enriched in the IP vs Input, but there were still many reads within the RNA in the IP sample, compared to negative control RNA, *GAPDH*, which had very few reads in the IP sample.

RNAs bound by TRIM71. Furthermore, most of the notable RNAs from the TRIM71 CLIP that we've already mentioned are actually bound by IMP1 as well, including *CDKN1A*, *EGR1*, *LIN28B*, *MDM2*, the *MBNL* family, and *NFIB*. Next we tested whether the overlapping RNAs also bound by IMP1 were more enriched in the TRIM71 CLIP, but in fact they were slightly less enriched than the 1721 non-IMP1 targets (Figure 2.8B). However, the IMP1 bound RNAs did represent a higher percentage of T to C transitions on average than non-IMP1 targets (Figure 2.8C)

A particularly notable IMP1 target is *MYC* due to its role as an oncogenic transcription factor and importance in development. In fact, an alias for IMP1 is CRDBP or coding region determinant-binding protein, so named due to the binding of IMP1 to the CRD element of the *MYC* mRNA. The CRD element of *MYC* is a highly structured hairpin element towards the 3' end of the coding sequence that contains a high number of rare codons (Lemm and Ross, 2002). The high level of structure in addition to the rare codons result in ribosomal pausing at this site within the RNA. The subsequent depletion of ribosomes towards the 3' end of the mRNA facilitates faster degradation of an already unstable transcript. The half-life of *MYC* is around 30 minutes. The *MYC* mRNA can be stabilized by the binding of IMP1 to this element or by efficient translation with minimal pausing (Lemm and Ross, 2002). Although this is a canonical role for IMP1 in the regulation of RNAs, *MYC* does not always come out as a hit in CLIP experiments. In fact, it is often weakly bound showing low enrichment of less than a 0.4 fold increase in a CLIP from ES cells (Conway et al., 2016). A meta-analysis of 8 different IMP1 CLIP datasets did not find *MYC* to be enriched as a target RNA (Müller et al., 2018). This begs the question, if IMP1 has been experimentally validated to bind *MYC* by multiple independent groups, why does it rarely come up as a hit in CLIP data sets? It could be that the short

half-life of the RNA prevents IMP1 from binding efficiently and therefore it only turns up as being weakly bound.

MYC did not turn up as an enriched RNA target in the TRIM71 CLIP. The IP samples actually show a slight depletion of the RNA by 0.7-fold in the IP compared to input in the differential expression analysis. The TRIM71 CLIP did still have many reads in the IP however there were less than the input samples (Figure 2.8D). We compared binding and enrichment within the *MYC* mRNA to the *GAPDH* mRNA and found that there were comparatively more reads in the *MYC* mRNA than *GAPDH*. As a highly expressed mRNA, *GAPDH* often turns up as enriched in CLIP experiments (Van Nostrand et al., 2016). The depletion of *GAPDH* here suggests that our washes were harsh enough to deplete for non-specific RNA binding. This agrees with previous results showing that highly expressed RNAs were not more strongly enriched in the IP samples (Figure 2.5B). As previously mentioned, PAR-CLIP causes predictable mutations where a protein is actually bound to the RNA. Looking at the *MYC* mRNA we do see the presence of thymine to cytosine transitions (Figure 2.5F lowest magenta point). In fact, *MYC* has a higher percentage of transitions than known targets *EGR1* or *CCNE2*. Due to the unique challenges of finding the *MYC* mRNA as a target in CLIP experiments we decided to follow up on it nonetheless.

TRIM71 regulation of target mRNAs

CLIP-qPCR

To follow up on the potential target RNAs we performed a CLIP followed by qPCR. The CLIP was performed in HA-tagged *TRIM71* expressing cells as well as empty vector

cells giving us a control IP sample to compare. Because the CLIP-seq samples were not compared to a control IP sample we only saw strongly bound genes that were enriched in the IP vs input. Using a control IP with the CLIP-qPCR allowed us to detect RNAs that were weakly bound. As a control we used an empty vector, allowing us to use the same HA antibody for both the control and TRIM71 IP. During an IP, RNAs are sometimes able to bind non-specifically to either the beads or the antibodies so using the same antibody for both conditions controlled for this factor.

With CLIP-qPCR we confirmed that TRIM71 was in fact binding to *CDKN1A*, *MBNL1/2/3*, *NFIA/X*, *LIN28B*, *MDM2*, *TCF7*, *TCF7L2,* and *MYC* (Figure 2.9A). IP samples of both the empty vector and the TRIM71 IP were first normalized to input then compared to each other for statistical significance. *ACTIN* and *GAPDH* were not significantly enriched in the IP indicating that TRIM71 was not binding these mRNAs. *EGR1* and *NFIB* were also not enriched in the CLIP-qPCR despite being enriched in the CLIP-seq. *EGR1* and *NFIB* are lowly expressed in HepG2 cells leading to large differences between the biological replicates. Of the mRNAs tested, *CDKN1A* was enriched the most in the TRIM71 CLIP-qPCR supporting that it was also a top hit in the CLIP-seq. *MYC*, despite not being enriched in the CLIP-seq was enriched in the CLIP-qPCR. However, its enrichment is noticeably less than the other target mRNAs which could explain its absence in the CLIP-seq IP vs input differential expression.

Next, we wanted to determine where in the target RNAs TRIM71 was binding. Following up on two target RNAs, *MYC* and *LIN28B*, we designed primers to tile along the mRNAs. For *MYC* we had 9 different primer sets: 2 within the 5’ UTR, 5 within the coding region, and 2 within the 3 ’UTR. Of these, primer set 8 was within the CRD element and set 7 partially overlapped with the 5’ end on the CRD element (Figure 2.9B). All

Figure 2.9. TRIM71 CLIP-qPCR.
(A) Fold enrichment of mRNAs in CLIP of cells expressing *HA-TRIM71* versus empty vector control. Known TRIM71 target *CDKN1A* and novel targets MBNL1/2/3, NFIA/X, *LIN28B*, *MDM2, TCF7L2, and MYC* were significantly enriched in the TRIM71 CLIP whereas control genes *ACTIN* and *GAPDH* were not. NFIB was not significantly enriched with a p-value of 0.055. **(B)** Performing CLIP qPCR with primer pairs tiled across the *MYC* mRNA revealed TRIM71 enrichment within the CRD element, the location where IMP1 is known to bind. **(C)** Primer pairs tiled across the *LIN28B* mRNA showed enrichment within the CDS and 3'UTR, consistent with where IMP1 is known to bind. All error bars represent standard deviation of fold change based on 3 or more biological replicates. Significance was determined by Welch's *t*-test with * and ** representing p values less than 0.05 and 0.01, respectively.

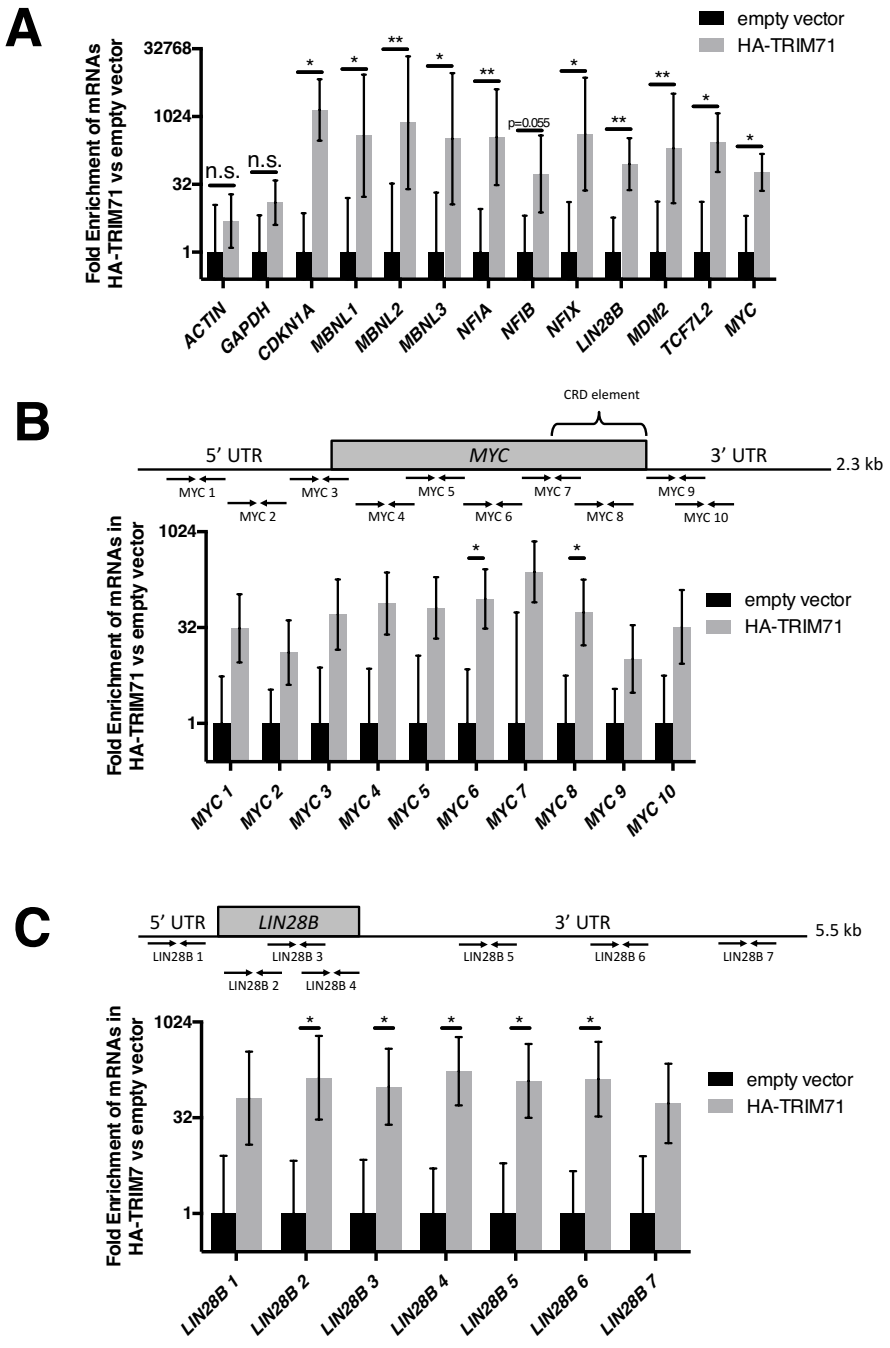


Figure 2.9 (Continued)

primers had equal amplifying efficiency so any differences in amplification should directly relate to the amount of template present. By CLIP-qPCR we observed significant binding to the mRNA around primer set 6 and set 8 which is in the CRD element. We also observed slightly increased enrichment of primer sets towards the middle of the mRNA compared to the 5' and 3' end which is likely due to instability of the *MYC* mRNA and degradation by exonucleases. Often in CLIP experiments, samples are treated with RNase to narrow the binding region to a discrete area, as the protein will protect this area from degradation. Although these samples were not treated with RNase, the instability of the *MYC* mRNA may be having the same effect which could explain the narrow area where we see TRIM71 binding.

Following up on *LIN28B* we designed 7 primer sets along the mRNA: 1 in the 5' UTR, 3 within the coding region, and 3 within the long 3'UTR (Figure 2.9C). Here we see enrichment of binding throughout the RNA with the exception of the most 5' and 3' ends. IMP1 is known to bind within the 3' UTR so this data appears consistent with that. As *LIN28B* is a more stable RNA than *MYC*, we do not see narrow enrichment where TRIM71 is binding. Given the broad enrichment of TRIM71 along the *LIN28B* mRNA, this could imply a lack of specify. However, to fully test this we would need to repeat the experiment in the presence of RNase in order to digest away the regions of the RNA where TRIM71 is actually not binding. In the absence of RNase, TRIM71 binding to a stable RNA such as LIN28B would likely pull down the entire RNA and we may not see site-specific enrichment.

TRIM71 differentially affects targets at the RNA and protein levels

CDKN1A and *EGR1* are both previously discovered targets of TRIM71 that were shown to be inhibited. To confirm this finding in HepG2 cells we knocked down *TRIM71*

by siRNA and collected protein lysates. We confirmed good TRIM71 knockdown and observed increases in p21, the protein encoded by *CDKN1A*, and EGR1 (Figure 2.10A, C). Interestingly, although *CDKN1A* is a strong TRIM71 target, the effect at the protein level was fairly small when quantified (Figure 2.10C). We repeated this experiment in Huh6 cells, a hepatoblastoma cell line, seeing an increase in p21 with *TRIM71* knockdown (Figure 2.10A, D).

After confirming that TRIM71 inhibits the known target mRNAs we then tested its effect on the new targets. For the *MBNL* and *NFI* families we expected TRIM71 to inhibit them, similar to the known targets. We tested MBNL1 and MBNL3 and did in fact see increased protein levels of MBNL1 with *TRIM71* depletion (Figure 2.10A, C). There was an inconsistent effect on MBNL3. We did not see a change in NFIA at the protein level despite seeing strong enrichment in the IP and high numbers of T to C transitions. This could indicate that *NFIA* is not regulated in this cell line, despite being bound, or perhaps it's not regulated at all.

Since we have shown that TRIM71 binds to IMP1, and knowing that IMP1 tends to promote RNAs bound, we wanted to test whether TRIM71 could do the same. First, we looked at whether IMP1 was promoting *MYC* in HepG2 cells. We knocked down *IMP1* with siRNAs and saw a decrease in MYC at the protein level confirming this canonical role for IMP1 (Figure 2.10B, E). Next, we depleted *TRIM71* and looked at the novel targets LIN28B, MDM2, TCF7, TCF7L2, and MYC. Interestingly, *TRIM71* knockdown decreased these targets at the protein levels suggesting that TRIM71 was promoting their RNAs. We also repeated these observations in Huh6 cells. We predicted that TRIM71 would promote these targets based on their depletion in the *Trim71* knockout mouse ES cells and we confirmed that prediction here in HepG2 and Huh6 cells (Figure 2.10B, C).

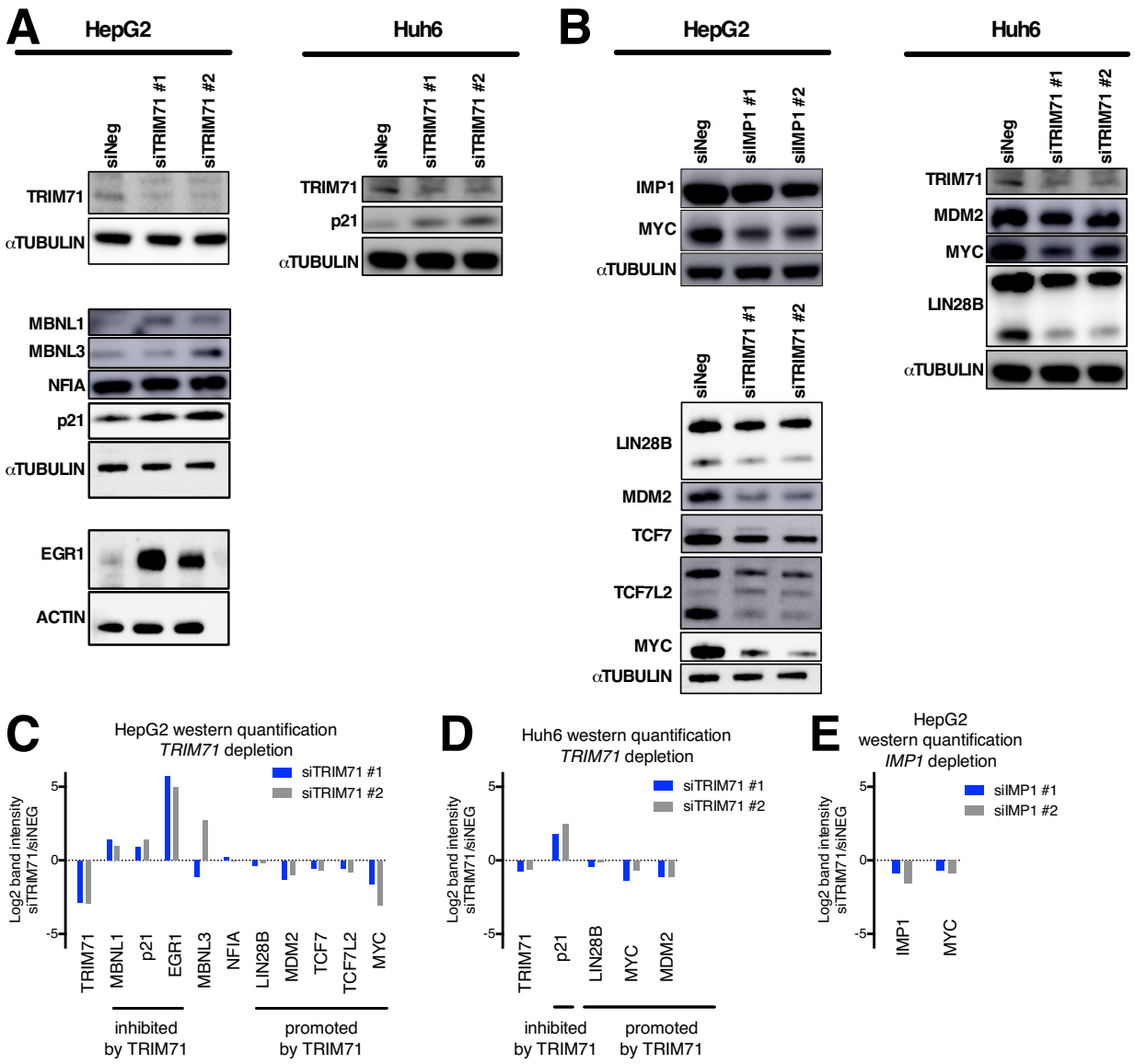


Figure 2.10. TRIM71 affects protein levels of known and novel targets in HepG2 and Huh6 cells.

(A) Known TRIM71 targets p21(*CDKN1A)* and *EGR1* were negatively regulated by TRIM71 and knockdown of *TRIM71* increased their expression at the protein level. Novel targets MBNL1/3 were also increased in TRIM71 knockdown. NFIA was not affected. **(B)** Knocking down *IMP1* decreased levels of MYC. Likewise, knockdown of *TRIM71* also decreased levels of MYC indicating TRIM71 positively regulates it. LIN28B, MDM2, TCF7, and TCF7L2 are also decreased upon *TRIM71* knockdown. **(C, D, E)** Western quantification of target proteins. Signal of blots from A, B were quantified then normalized to the loading control. It was normalized again to siNEG and log2 differences were plotted. Positive values indicate increase levels with *TRIM71* depletion whereas negative values indicate decresed levels. Replicate westerns showed similar enrichment and de-enrichment patterns.

After seeing an effect at the protein level with *TRIM71* knockdown, we next looked at the RNAs. *TRIM71* knockdown decreased the levels of target RNAs *MDM2* and *LIN28B* but not *MYC* (Figure 2.11A). This suggests that TRIM71 could be promoting the stability of *MDM2* and *LIN28B*. A similar result was observed in Huh6 cells for *LIN28B* and *MYC* (Figure 2.11B). An analysis of CRD complex proteins on the stability of the *MYC* RNA has shown that only some of the complex members directly promote the stability of the RNA, while others like STAU1 and HNRNPE2 do not affect decay rates of *MYC* (Weidensdorfer et al., 2009). However, *STAU1* knockdown still regulated post-transcriptionally, decreasing the production of the MYC protein, presumably by affecting translation of the mRNA. Interestingly, STAU1 and HNRNPE2 are not always associated with the CRD complex. Their association may be cell type dependent. This shows that there are complex members that do not directly affect mRNA stability which could account for the lack of effect of *TRIM71* depletion on the *MYC* mRNA.

While measuring the steady state levels of mRNAs as in Figure 2.11A/B, it is not possible to determine whether the decrease in the *MDM2* and *LIN28B* RNAs was due to transcriptional or post-transcriptional regulation. To separate these two possibilities, we performed a luciferase reporter experiment with the 3' UTR of *LIN28B* cloned behind a renilla luciferase gene driven by a T7 promoter. Based on the CLIP-seq and CLIP-qPCR we thought TRIM71 was binding within the *LIN28B* 3' UTR. Any changes in reporter activity would relate directly to the *LIN28B* 3' UTR sequence and not differences in transcription. The luciferase reporter showed less activity when *TRIM71* was depleted by siRNAs (Figure 2.11C). This suggests that TRIM71 was mediating a post-transcriptional effect on *LIN28B*, possibly by directly affecting stability of the RNA.

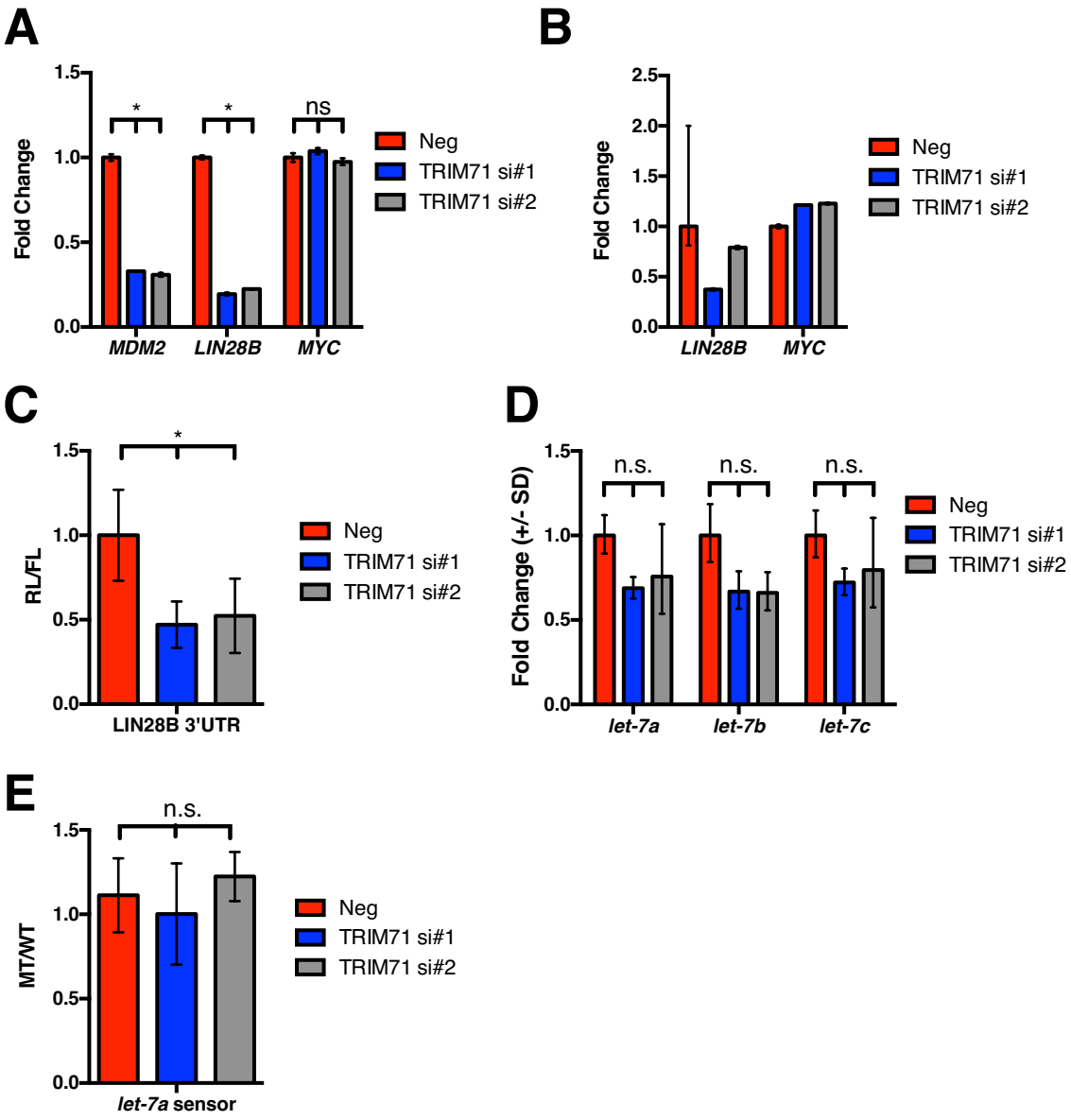


Figure 2.11. TRIM71 affects mRNA levels of LIN28B and MDM2 but not MYC. TRIM71 does not affect *let-7* levels or activity.

(A) *TRIM71* knockdown decreased mRNA levels of MDM2 and *LIN28B* but not *MYC* in HepG2 cells. Error bars represent S.D. of two (MDM2) or three (LIN28B, MYC) biological replicates. Significance was determined by *t*-test $p<0.05$. **(B)** TRIM71 knockdown decreased *LIN28B* mRNA but not *MYC* in Huh6 cells. Error bars represent S.D. of two biological replicates. **(C)** A PsiCheck2 luciferase reporter with the *LIN28B* 3'UTR showed decreased activity with *TRIM71* knockdown. Error bars represent S.D. One-way ANOVA, $p<0.05$. **(D)** *TRIM71* depletion did not affect *let-7* levels measured by qPCR. One-way ANOVA. **(E)** A PsiCheck2 *let-7a* sensor was not changed with *TRIM71* knockdown. MT refers to a non-complementary mutant *let-7a* sensor while WT is fully complementary. Not significant based on one-way ANOVA.

TRIM71 does not affect activity of let-7

Since many of the TRIM71 RNA targets are themselves targets of *let-7,* we wanted to rule out the possibility that *let-7* was mediating any changes we saw by post-transcriptional means, as it's possible that the *TRIM71* RNA could be directly impacting levels or activity of *let-7* by a sponging effect. Knockdown of *TRIM71* did not significantly change *let-7* levels as measured by qPCR (Figure 2.11D). Had a feedback effect on *let-7* been responsible for the decrease in *LIN28B*, we'd actually expect an increase in *let-7* levels or activity as opposed to the insignificant decrease we saw.

Next we performed a luciferase reporter assay with the *let-7a* complementary sequence cloned into the 3' UTR of renilla luciferase and compared activity to a reporter with a mutated sequence. Knockdown of *TRIM71* did not affect activity of the *let-7* reporter which likely rules out the possibility that the *TRIM71* RNA is acting as a *let-7* sponge (Figure 2.11E). Therefore, TRIM71 does not appear to affecting the *LIN28B* mRNA levels through a feedback effect on *let-7*.

IMP1 helps specify where TRIM71 binds

With the understanding that TRIM71 and IMP1 are binding RNAs as members of a complex, we wanted to better understand the influence they had on each other with regards to RNA binding. We knocked down *IMP1* by siRNA and tested the effect of TRIM71 binding to its target RNAs. With *IMP1* depleted TRIM71 was able to bind *ACTIN* significantly more strongly. This indicated decreased specificity of TRIM71 binding (Figure 2.12A). Looking at TRIM71 targets, in general we saw a small increase in binding, though only MYC was significantly more enriched. This is surprising since IMP1 is known to bind *MYC* and it seems counterintuitive that a known *MYC* binding protein would inhibit

TRIM71 binding to *MYC* if they are a part of the same complex. It is possible that TRIM71 on its own is not binding RNA with as much specificity as the complex as a whole. Since the complex is composed of multiple RNA binding proteins, each with their own recognition motif, the combination of two or more motifs would mean there are fewer binding spots that the complex as a whole will recognize compared to any individual component. Together with the binding to *ACTIN* this suggests that TRIM71 binds promiscuously with *IMP1* depleted and binds to some weak targets more effectively, though the effects are generally small.

Next, we tested what impact TRIM71 had on the RNA binding of IMP1. Knockdown of *TRIM71* did not affect IMP1 binding to any RNAs tested (Figure 2.12B). IMP1, as a key member of the CRD complex, is likely still able to form the complex without TRIM71. This may explain why *TRIM71* depletion does not affect IMP1 binding to RNA. On the other hand, depletion of IMP1 may dissociate the complex completely, thereby affecting how TRIM71 binds to RNA as shown above.

RNA binding of TRIM71 domain deletion mutants

The NHL domain on its own binds RNA with less specificity

To better understand how TRIM71 binds to RNAs on its own, without influence of other RNA binding proteins, we performed a CLIP-seq of just the RNA-binding NHL domain. We previously showed that the NHL domain did not bind to CRD complex members IMP1 or DHX9 whereas other domain deletion mutants did still bind IMP1 (Figure 2.3). A comparison of the NHL CLIP-seq with a wild-type TRIM71 CLIP would allow us to determine what impact protein binding partners have on TRIM71 RNA binding.

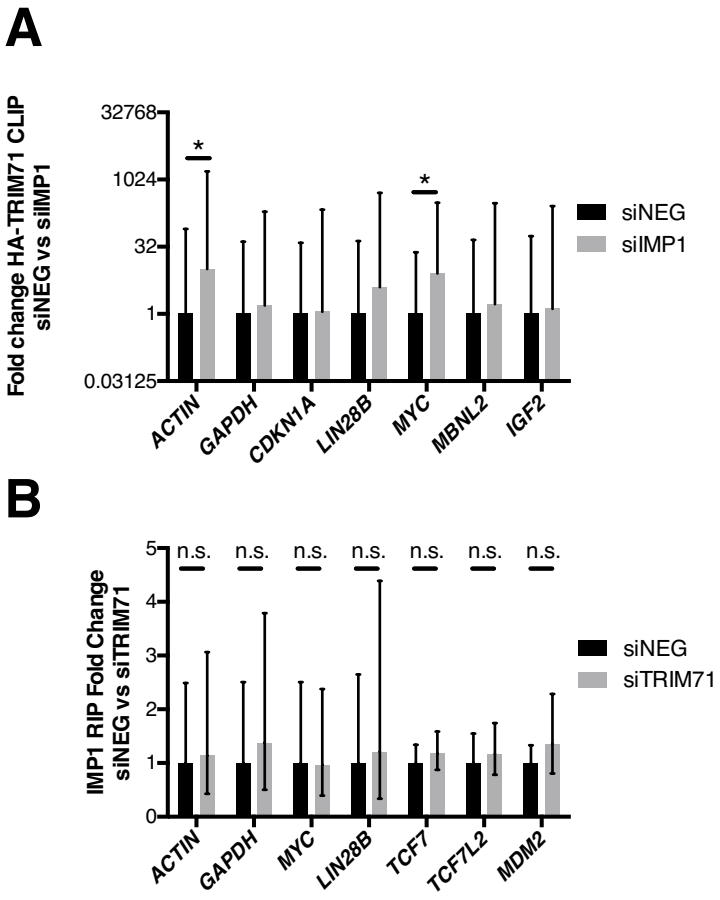


Figure 2.12. IMP1 may specify TRIM71 binding while TRIM71 does not affect IMP1 binding.

(A) CLIP of TRIM71 following siRNA treatment against IMP1 or a non-targeting control. IMP1 knockdown increased binding affinity of TRIM71 to some targets and decreased binding specificity. Genes indicated were significantly enriched by *t*-test $p<0.05$. Genes not indicated were not enriched. **(B)** Knockdown of TRIM71 did not affect IMP1 binding to any known targets tested. Error bars show S.D. of 3 biological replicates. Significance tested by t-test.

We performed a differential expression analysis of the NHL CLIP IP with the NHL input finding 2333 RNAs enriched in the NHL IP. This is a similar number as the 2432 RNAs that were enriched in the TRIM71 IP versus input. We plotted the NHL differential expression analysis against the wild-type differential expression to compare all RNAs, not just those enriched in the IPs. There were 2362 genes that were more than 2-fold enriched in the NHL IP compared to the TRIM71 wild-type IP (Figure 2.13A). Comparatively, there were only 1790 genes more than 2-fold enriched in the WT versus the NHL domain. This suggests that in general the NHL domain is binding RNAs more strongly than wild-type TRIM71 or with less specificity. To further compare, we took the difference of enrichment in NHL versus wild-type for each gene and plotted a histogram of the differences. Genes more strongly bound by NHL had a fold change difference greater than 0 while those bound more strongly by wild-type TRIM71 had a difference less than 0. The distribution of the histogram was significantly shifted in the positive direction indicating that RNAs tend to be more strongly bound by the NHL domain on its own compared to wild-type TRIM71 (Figure 2.13B).

Focusing on genes depleted in the wild-type TRIM71 IP, 1679 were enriched in the NHL IP (Figure 2.13C) indicating that the NHL domain is able to bind more RNAs that full-length TRIM71 does not normally bind. Looking further at the T to C transitions of these 1679 NHL target RNAs, they represent a higher percentage of transitions in the NHL CLIP compared to the full-length TRIM71 CLIP (Figure 2.13D) indicating that these RNAs are truly bound by the NHL domain on its own. Next, we looked at T to C transition percentages of the 2432 TRIM71 target RNAs and saw fewer transitions in the NHL-only CLIP (Figure 2.13E). Overall, this indicates that TRIM71 target RNAs are bound less strongly by the NHL domain on its own but the NHL domain binds with more promiscuity.

Figure 2.13. The NHL domain on its own binds promiscuously compared to full-length TRIM71.
(A) Scatter plot of differential expression scores of genes in TRIM71 IP vs input plotted against NHL-only IP vs input. Genes above or below the blue dotted lines were 2 fold over expressed or 2 fold depleted in the NHL-only IP compared to full-length TRIM71 IP. There were 2362 genes overexpressed and 1790 genes depleted in the NHL-only IP compared to full-length TRIM71. **(B)** Histogram representation of genes in subfigure A calculated by taking the difference between the NHL-only enrichment and full-length TRIM71 enrichment. The histogram distribution is significantly shifted from zero in the direction of enriched in NHL-only, $p<0.0001$, one sample *t*-test. **(C)** Of the genes depleted in full-length TRIM71 IP, 1679 of them were enriched in the NHL-only IP. **(D)** T to C conversion fractions of the 1679 genes bound by NHL but not by TRIM71 show more conversions in the NHL IP versus the TRIM71 IP, indicating a true hit. Paired *t*-test, $p=0.0001$. **(E)** T to C conversion fractions of the 2432 TRIM71 hits showed fewer conversions in the NHL CLIP versus the TRIM71 CLIP indicating less binding to targets. Paired *t*-test, $p<0.0001$ **(F)** Testing TRIM71 binding at the individual gene level by CLIP followed by qRT-PCR showed slight enrichment of non-target genes *ACTIN* and *GAPDH*, though only *GAPDH* was significantly enriched. Target genes were not changed.

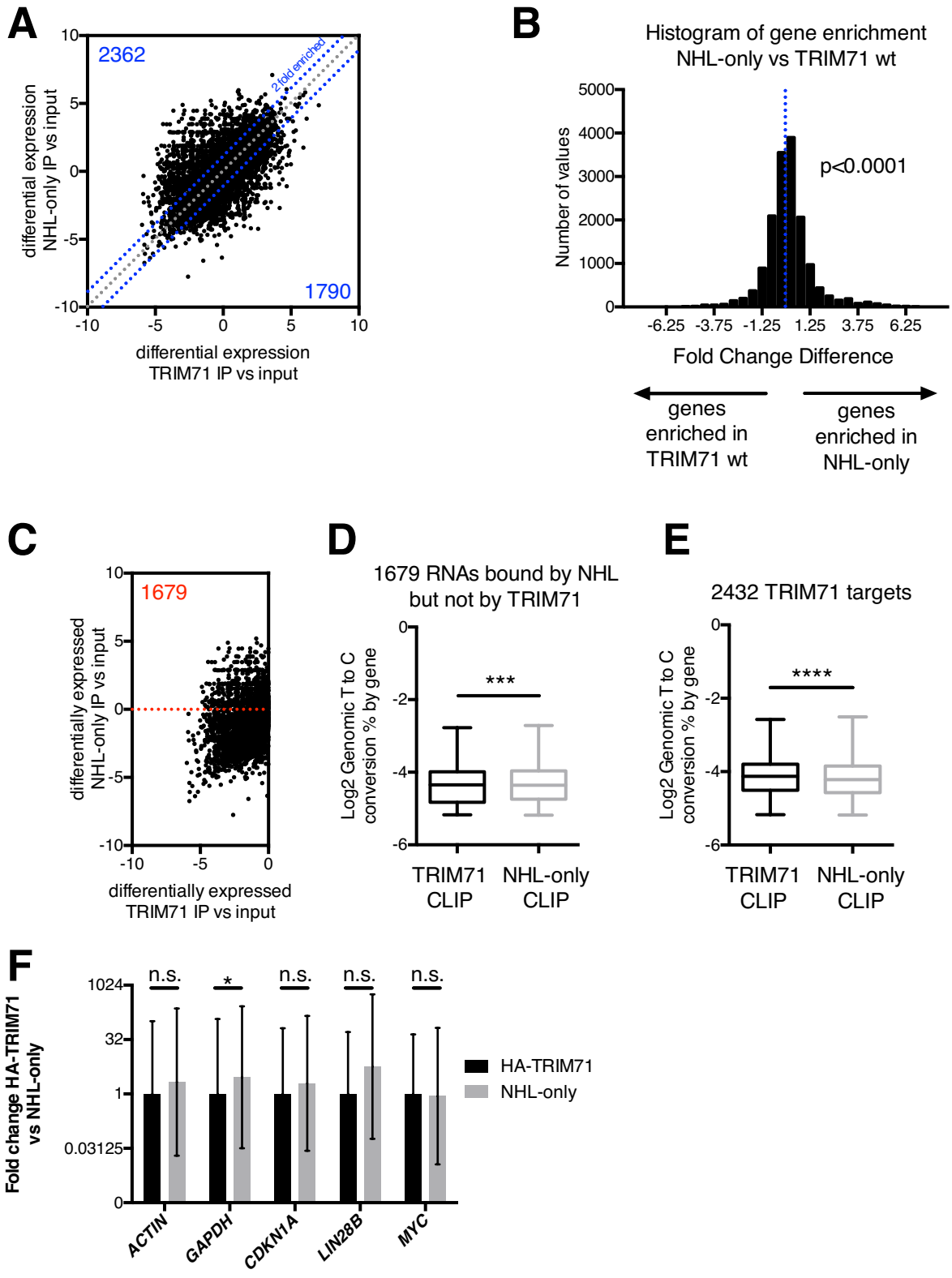


Figure 2.13. (Continued)

To confirm the NHL CLIP-seq results on an individual gene level we performed a CLIP-qPCR of the NHL domain and showed a small but significant increase in binding to the *GAPDH* control gene which is not normally bound by TRIM71 (Figure 2.13D). TRIM71 targets *CDKN1A*, *LIN28B*, and *MYC* were not changed (Figure 2.9A). Based on depletion of *IMP1* we hypothesized that TRIM71 binding to RNAs was promiscuous on its own. The data here of the NHL domain CLIP-Seq supports that hypothesis.

Deletion of the NHL domain prevents TRIM71 binding to most target mRNAs

We performed a CLIP-seq of an NHL deletion mutant TRIM71, ΔNHL, and analyzed the differential expression of RNAs in the IP versus input to understand what RNAs were still enriched when the RNA binding domain of TRIM71 was deleted. We plotted the enrichment values of the 2432 TRIM71 target RNAs against their enrichment in the ΔNHL IP (Figure 2.14A). Many of the genes appear below zero on the y-axis and below the dotted blue lines which signifies a 2-fold depletion, indicating that they are less enriched in the ΔNHL CLIP.

To quantify this effect, we took the difference in the fold change enrichment of the ΔNHL and TRIM71 wild-type CLIPs. Scores greater than zero indicate genes that are more enriched in the ΔNHL IP whereas scores less than zero indicate genes that are depleted in the ΔNHL IP and enriched in the full-length TRIM71 IP. The distribution of scores is significantly shifted in the negative direction indicating that the majority of TRIM71 target genes are depleted in the ΔNHL IP (Figure 2.14B). As expected, deleting the RNA binding domain of TRIM71 inhibits its interaction with target RNAs.

To confirm this, we performed a CLIP-qPCR to look at top TRIM71 targets. Target RNAs *CDKN1A*, *MYC*, and *LIN28B* were bound less strongly (Figure 2.14C). The non-

target mRNAs of *ACTIN* and *GAPDH* did not decrease in the ΔNHL CLIP-qPCR as they were not bound to begin with.

Counterintuitively, some targets were bound more strongly by the ΔNHL mutant. One of the more enriched targets in the ΔNHL samples was *IGF2*, which had nearly 10x the number of reads in the ΔNHL CLIP samples compared to wild-type TRIM71 CLIP (Figure 2.14D). *IGF2* is a known RNA target of IMP1. In fact, an alias for IMP1 is IGF2 RNA binding-protein or IGF2BP1. To understand why the ΔNHL mutant is able to bind this RNA we have to look back at its coprecipitated binding partners (Figure 2.3B). The ΔNHL mutant is still able to coprecipitate with IMP1 and DHX9 indicating that the CRD complex is likely still intact and TRIM71 is binding it. However, without the RNA binding domain of TRIM71, any RNAs actually bound would be specified by the other members in the complex with their own RNA binding domains, such as IMP1 and DHX9. As shown earlier, around ~30% of TRIM71 target RNAs were also discovered in published IMP1 CLIP studies (Figure 2.8A) (Conway et al., 2016). Therefore, we wanted to test whether known IMP1 targets were bound differently from non-IMP1 targets in the ΔNHL CLIP. We split the genes from Figure 2.14B and replotted the scores based on IMP1 targets and non-IMP1 targets. Although both groups still showed depletion in the ΔNHL IP, not surprisingly, the IMP1 targets were slightly less depleted (Figure 2.14E).

The NHL domain has an important impact on how TRIM71 binds RNAs. RNAs are significantly less enriched in ΔNHL CLIP samples compared to wild-type TRIM71. However, because the ΔNHL mutant TRIM71 is able to bind other RNA binding proteins, we can consider this mutant as still having the capacity to bind RNA based on the affinities

Figure 2.14. Deletion of the NHL domain prevents TRIM71 from binding to most target mRNAs.
(A) Differential expression of TRIM71 targets plotted again enrichment in ΔNHL CLIP. **(B)** Fold change differences show that genes are significantly less enriched in ΔNHL compared to full-length TRIM71. **(C)** CLIP followed by qRT-PCR showed that deletion of the RNA-binding NHL domain decreased the affinity of TRIM71 for target RNAs tested. **(D)** Deletion of the NHL domain of TRIM71 causes it to bind more strongly to IGF2 as shown by nearly 10x more reads in the ΔNHL samples (blue) compared to full-length TRIM71 (black). **(E)** Fold change differences of IMP1 vs non-IMP1 target RNAs showed that IMP1 targets were significantly more bound than non-IMP1 targets in the ΔNHL CLIP.

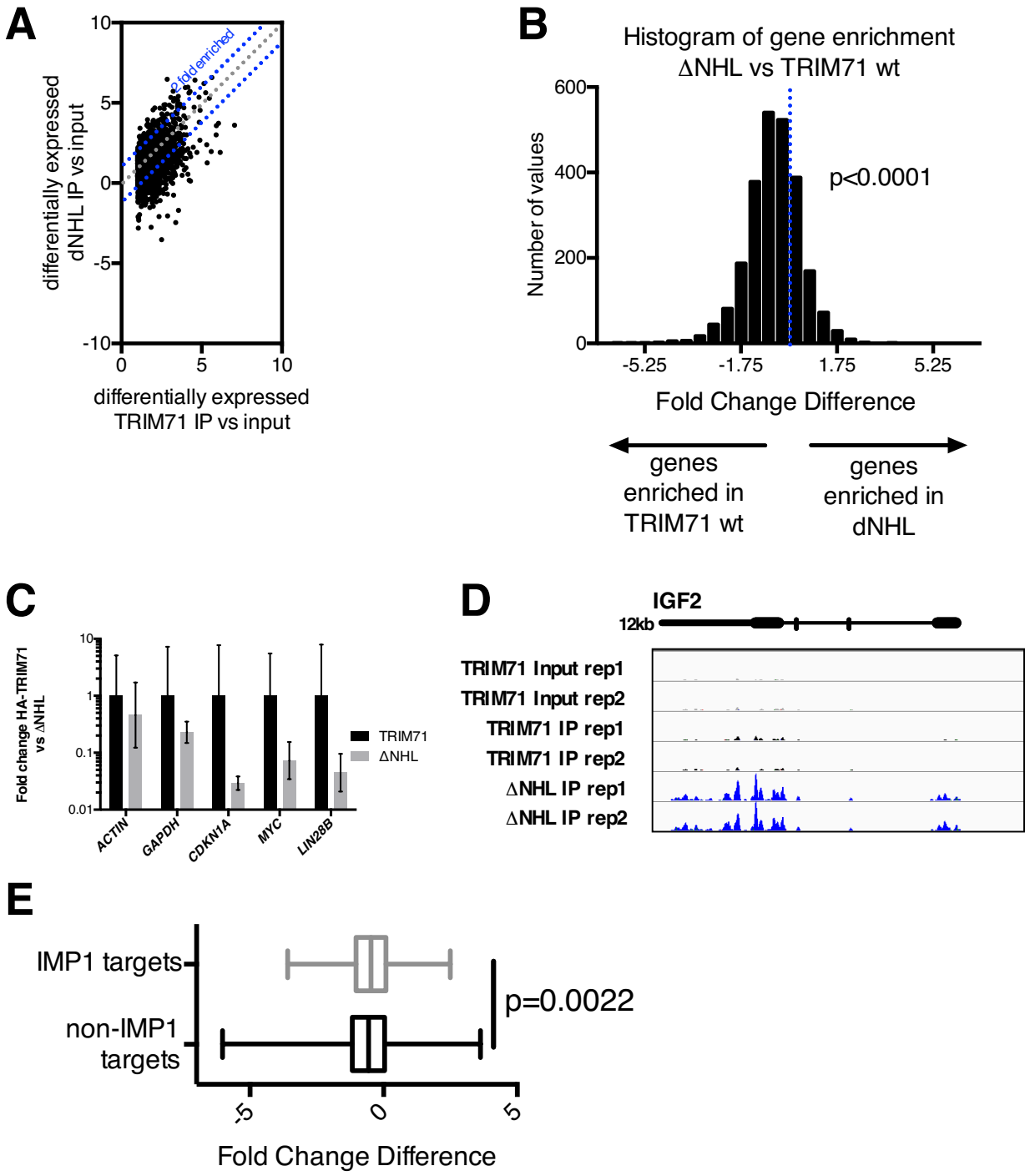


Figure 2.14 (Continued)

of proteins such as IMP1. It is unlikely that the other TRIM71 domains are binding RNA as none have been described to do so.

NHL point mutations decrease enrichment of TRIM71 targets and bind to new RNAs

In addition to performing CLIP of wild-type TRIM71 and the NHL domain on its own, we also had two TRIM71 mutants with point mutations in the NHL domain. These points mutations, R608H and R796H, were the residues mutated in congenital hydrocephalus patients and were located on the positively charged RNA binding surface of the NHL domain (Figure 1.3B). So, we reasoned that RNA binding was likely affected by these mutations, leading to the patient phenotypes. Arginine (R) is more polar than histidine (H) so the RNA binding is probably not the same. However, because arginine and histidine are both positively charged amino acids these mutants may still be able to bind RNA, thus we performed the CLIP to find out.

CLIP replicates for both NHL point mutants were well correlated with Pearson coefficients $r>0.94$ (Figure 2.15A). Looking at sample clustering based on correlations of transcript reads, the IP samples clustered away from the input (Figure 2.15B). Looking more closely at the IP samples, the NHL point mutants clustered closely with each other and with the ΔNHL mutant. Meanwhile the NHL-only domain clustered closely with wild-type TRIM71. The clustering of ΔNHL away from wild-type TRIM71 is expected due to the absence of the RNA binding domain. The NHL point mutants clustering with the ΔNHL mutant would indicate that these point mutations have a severe effect on RNA binding of TRIM71.

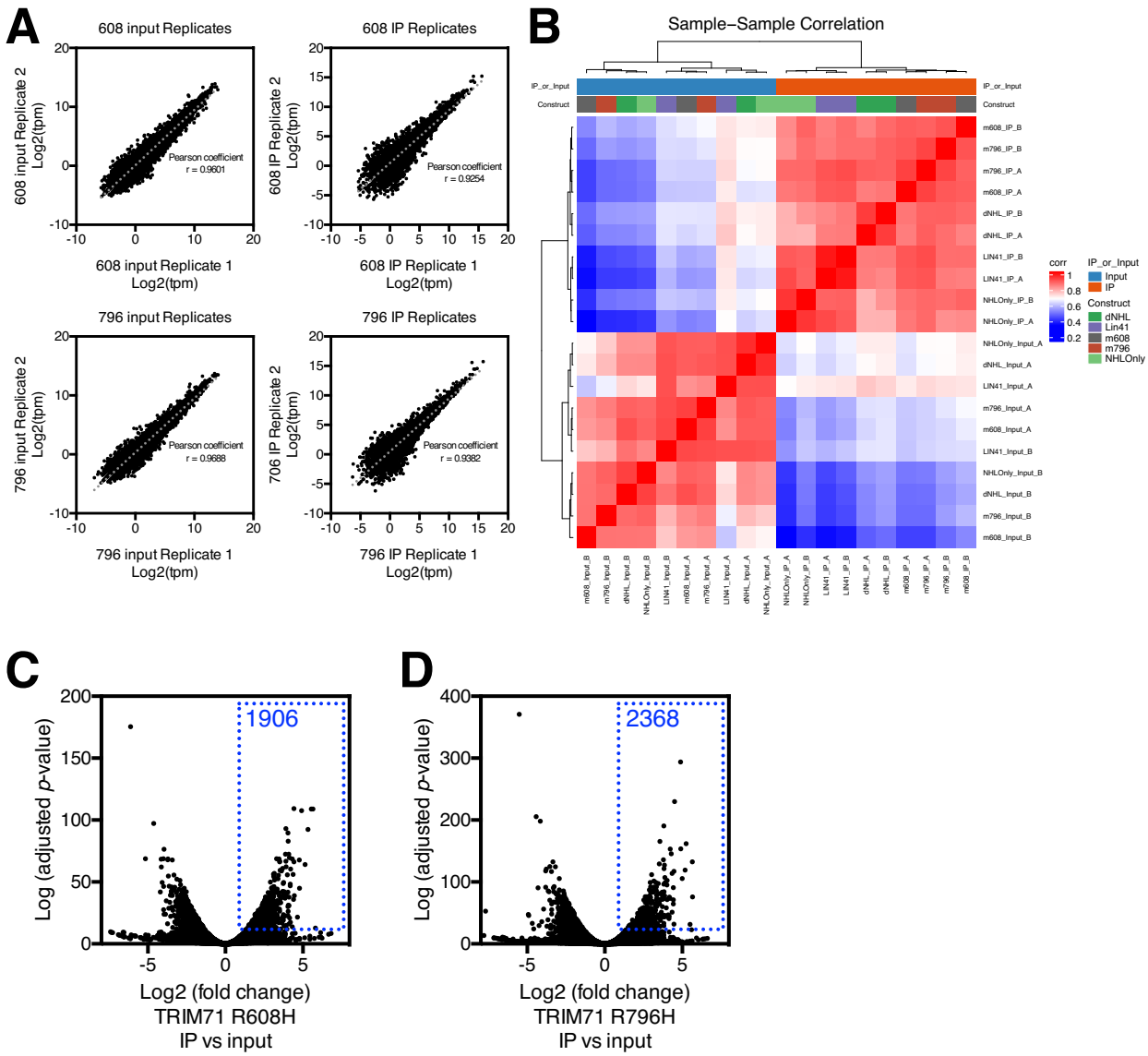


Figure 2.15. NHL point mutants cluster with dNHL mutant. However, they still bind a defined set of RNAs.

(A) CLIP sample input and IP replicates for the NHL point mutants were well correlated with Pearson coefficients greater than 0.93. **(B)** Sample-sample correlation which clusters samples based on similarities showed that all input samples correlated with each other and all IP samples correlated. Further, the TRIM71 wt clustered with NHL-only and the ΔNHL mutant clustered with the point mutants. **(C, D)** Volcano plot of IP vs input for the TRIM71 NHL point mutants showing that 1906 and 2368 RNA are significantly enriched in R608H and R796H IPs, respectively.

Next, we performed a differential expression analysis of the NHL point mutants comparing the IP samples to the matched inputs to understand what RNAs were enriched. Despite clustering with the ΔNHL mutant, each point mutant still bound a set of defined RNAs. The R608H mutant was significantly enriched for 1906 RNAs and the R796H mutant was enriched for 2368 RNAs greater than a fold change of 2 (Figure 2.15C, D).

Testing the possibility that these point mutations are loss of function, we looked at their ability to bind the 2432 RNA targets of wild-type TRIM71. We plotted the enrichment of these targets for each point mutant finding that many were still enriched (Figure 2.16A). A small subset of the RNAs bound by wild-type TRIM71 were actually depleted in both NHL point mutant IP samples (Figure 2.16A blue box). Despite seeing many TRIM71 target RNAs bound by the point mutants, we wanted to understand the strength of binding. Although we did not directly measure RNA affinity of the point mutants, we looked for differences in enrichment between the mutants and wild-type IP samples. Fold change difference scores less than 0 indicate RNAs enriched more strongly in wild-type TRIM71 samples while differences greater than 0 indicate stronger enrichment with the mutant. We plotted the score of each gene on a histogram and found that both mutants had distributions significantly less than 0 (Figure 2.16B). Therefore, a significant proportion of TRIM71 target RNAs, despite being bound by the NHL point mutants, were enriched less and possibly bound to a lesser degree. An alternative explanation is that the point mutations are actually causing a decrease in the specificity of RNA binding and increasing promiscuity. This could account for decreased enrichments of wild-type TRIM71 target RNAs if the NHL point mutants are binding other RNAs with much higher affinities through

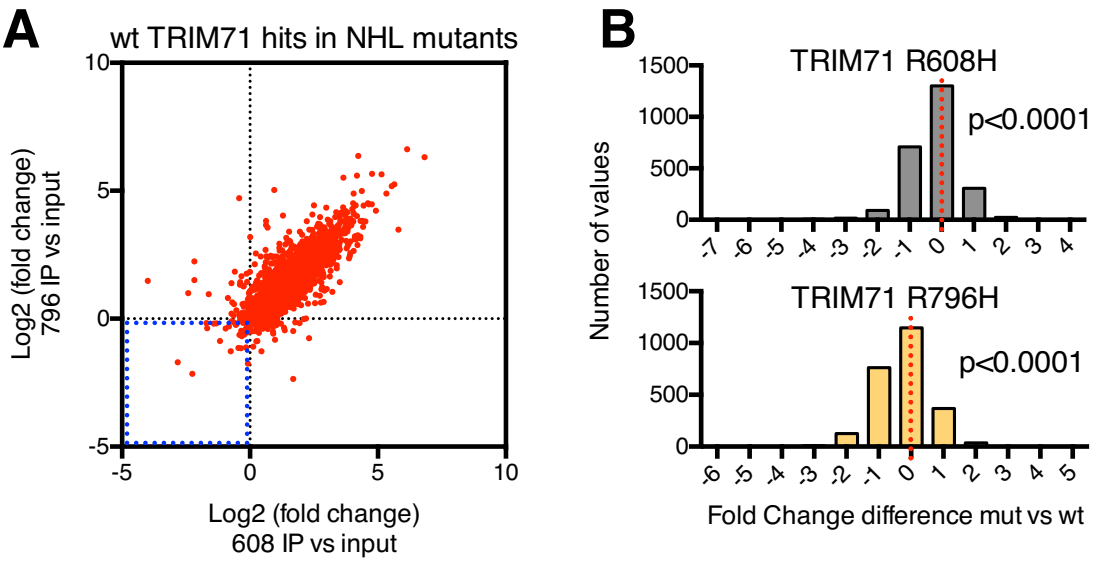


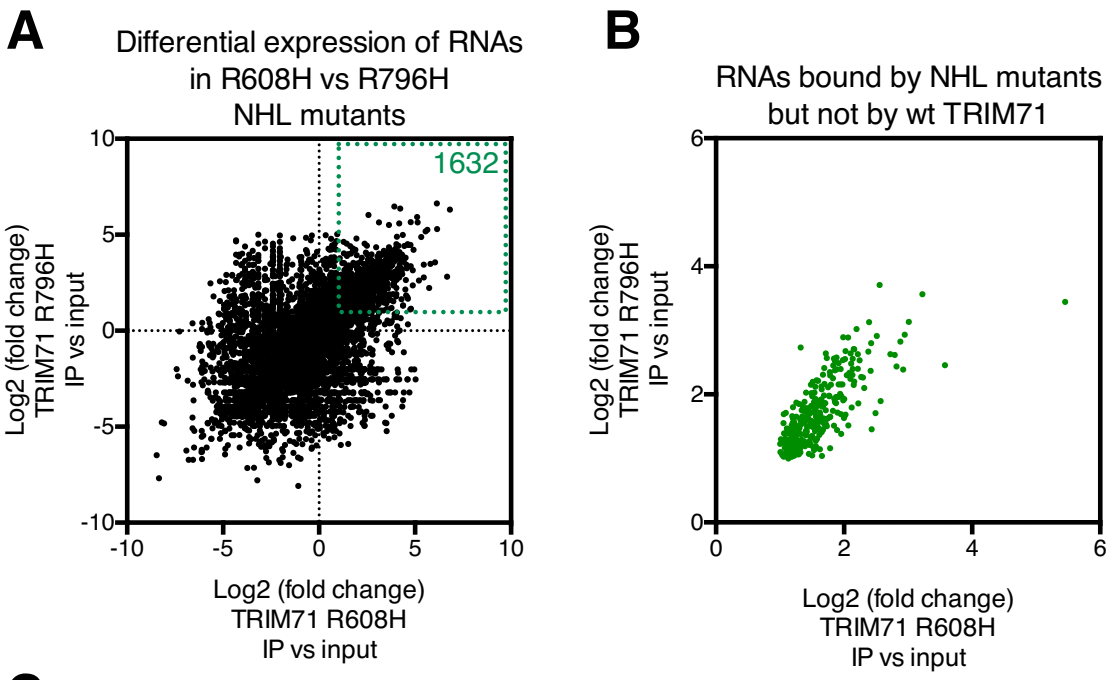
Figure 2.16. TRIM71 NHL point mutants show less enrichment for wild-type TRIM71 targets.

(A) Scatter plot of differential expression of wt TRIM71 targets in the NHL point mutants. Most RNAs were still significantly enriched in the IPs compared to inputs, upper-right quadrant. However, binding to several RNAs highlighted within the blue box was completely abrogated. **(B)** Histograms showing differences in the fold change enrichment of target RNAs in mutant IPs vs TRIM71 wt IP. Values less than zero correspond to RNAs bound less strongly by the mutant vs wt TRIM71. Both distributions had averages significantly less than 0, *t*-test, p<0.0001.

a gain-of-function effect.

Although it's possible that these NHL mutants cause a loss of function, it is also possible that they're causing a gain of function. These two mutations were observed in 4 congenital hydrocephalus patients heterozygous for TRIM71 mutations (Furey et al., 2018). The R608H mutation independently arose twice, with one of the patients passing it to a child, while the R796H mutation occurred in a homologous position of a different NHL repeat. For TRIM71, a gain of function could result from newly bound RNAs, not bound by the wild-type TRIM71. To test this, we looked for newly bound RNAs found in CLIPs of both mutants, reasoning that any gain of function in terms of bound RNAs would likely be shared between the two mutants. We found a total of 1632 RNAs significantly enriched by greater than a fold change of 2 in the CLIPs of both mutants (Figure 2.17A). We then filtered out RNAs that were bound by wild-type TRIM71 and were left with 307 newly bound RNAs that could represent a gain of function effect (Figure 2.17B).

We ran a gene ontology analysis on the newly bound mRNAs and found several metabolomic terms. Other GO terms include "poly-A ribonuclease activity, Ribosomal subunit export, and miRNA gene silencing" (Figure 2.17C). Nothing relating specifically to neural development arose in these newly bound RNAs. However, the caveat is that this CLIP was carried out in HepG2 cells which may not express the RNAs present during neural development and differentiation.



C

Molecular Function	Fold Change	FDR
poly(A)-specific ribonuclease activity (GO:0004535)	20.51	2.62E-02
NADH dehydrogenase (quinone) activity (GO:0050136)	8.69	2.80E-02
NADH dehydrogenase (ubiquinone) activity (GO:0008137)	8.69	2.66E-02
NADH dehydrogenase activity (GO:0003954)	8.51	2.82E-02
oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor (GO:0016655)	7.78	1.82E-02

Biological Function	Fold Change	FDR
ribosomal small subunit export from nucleus (GO:0000056)	28.57	4.10E-02
gene silencing by miRNA (GO:0035195)	8.33	2.19E-02
mitochondrial electron transport, NADH to ubiquinone (GO:0006120)	8.16	2.36E-02
mitochondrial ATP synthesis coupled electron transport (GO:0042775)	7.32	1.14E-03
ATP synthesis coupled electron transport (GO:0042773)	7.24	1.18E-03

Figure 2.17. NHL point mutants bind an overlapping set of RNAs.
(A) Scatter plot of differential expression of RNAs in the R698H and R796H NHL point mutants. There were 1632 RNAs significantly enriched in both point mutant IPs, highlighted by the green box. Among the 1632 RNAs, 307 of them were not bound by wt TRIM71. These are shown in **(B)**. **(C)** Gene ontology analysis of novel RNAs bound by NHL point mutants but not by wt TRIM71.

Summary

In this chapter we discussed the roles of TRIM71 as an RNA binding protein. We revealed that TRIM71 co-precipitates with members of the RNA-binding CRD complex, IMP1 and DHX9. A domain deletion analysis of TRIM71 showed that DHX9 interacts with the coiled-coil domain of TRIM71. However, we were not able to determine any one TRIM71 domain required for the interaction with IMP1. It is possible that multiple domains are mediating this interaction.

Imp1 is expressed at similar times and in similar tissues as *Trim71* in developing mice. Homozygous loss of *Imp1* leads to smaller adult mice and results in 50% perinatal lethality with these mice dying within 3 days of birth (Hansen et al., 2004). *Trim71* deletion leads to 100% embryonic lethality with mice dying before day E16.5. However, the *Trim71*-/- embryos are markedly smaller than wild-type *Trim71* mice. Although deletion of *Trim71* results in more severe phenotypes, the smaller body size caused by loss of either *Trim71* or *Imp1* could indicate similar regulatory effects in development. The most striking *Trim71*-/- phenotypes occur in the brain where the neural tube fails to close (Schulman et al., 2008). Mice with deleted *Imp1* do not exhibit the same strong phenotype but they have depleted neural progenitor cell populations within the dorsal telencephalon leading to smaller brain mass and a decline in adult neural stem cell functions (Degrauwe et al., 2016b).

The canonical role of IMP1 is to positively regulate RNAs. Given its binding to TRIM71 we challenged the idea the TRIM71 is a strict negative regulator of RNAs and tested whether it was capable of positive RNA regulation. Interestingly, IMP1 is capable of promoting or inhibiting RNA, so RNA binding proteins playing multiple roles is not unprecedented (Stöhr et al., 2012). After sequencing the RNAs bound by TRIM71 we

discovered that over half are likely to be positively regulated based on published *Trim71* knockout data. We confirmed that TRIM71 was in fact able to bind and positively regulate several RNA targets such as *LIN28B*, *MDM2*, *TCF7L2*, and *MYC.* The *Trim71* knockout mES cells were a good proxy for figuring out the direction of regulation for each individual RNA (Mitschka et al., 2015). With the use of luciferase reporters, we confirmed that the regulation of *LIN28B* was not transcriptional and not due to changes in *let*-7. Therefore, the positive regulation was most likely related to translation or stability of the RNA.

Next, we looked at the autogenous RNA binding of TRIM71 in two different ways to understand how the protein on its own binds without external influence. First, we depleted its binding partner IMP1 finding small changes in how certain RNA targets were bound. Second, we analyzed RNAs bound by just the NHL domain, which is not able to bind CRD complex members, finding that more RNAs were bound and full-length TRIM71 targets were enriched less. This suggests that TRIM71 binds RNAs more promiscuously on its own. Although TRIM71 is able to bind RNA differently on its own, it is still likely best understood in the context of complexed proteins. This is highlighted by the ΔNHL CLIP. Although deletion of the TRIM71 RNA binding domain had a strong impact on its ability to bind RNA, those targets shared with IMP1 were not depleted to the same degree. These CLIP sequencing experiments of TRIM71 domain deletion mutants suggest that the RNA binding activity of TRIM71 is intrinsically promiscuous. Thus, future studies looking into the RNA targets of TRIM71 should not exclude the impact of protein binding partners.

Lastly, we performed CLIP-seq on NHL point mutants to uncover differences that could result in patient phenotypes. This analysis uncovered global changes in how these

point mutants bind RNAs. These mutants showed less enrichment for wild-type TRIM71 targets, but also showed enrichment of several hundred new target RNAs. It is possible that the decreased enrichment of wild-type TRIM71 targets indicates weaker binding. However, it is also possible that higher affinity for new RNA targets could be sequestering the protein. Nonetheless, the end result is clear that there is a change in the RNAs bound.

The experiments discussed in this chapter provide a deeper understanding of the RNAs that TRIM71 binds, how different TRIM71 mutations affect this binding, and how protein binding partners influence interactions with RNA. Follow-up experiments should focus on the developmental impacts of these novel target RNAs in contexts such as neural progenitor cells or ES cells during neural differentiation. The *MBNL* and *NFI* families are both associated with neural development phenotypes when perturbed. As these are both potential targets of TRIM71, next steps should include analyzing the expression patterns of these RNAs when *TRIM71* is depleted in neural progenitors or ES cells. If these targets prove valid in the neural development context, follow-up experiments could include analysis of alternative splicing patterns caused by the *MBNL* family or transcriptional changes caused by the *NFI* family. Lastly, as the *MBNL* and *NFI* families are conserved from *C. elegans* to humans, it would be interesting to test whether *lin-41* perturbation in worms impacts expression of these targets and whether *lin-41* dependent phenotypes are related to expression of these novel targets.

Chapter 3 – *TRIM71* expression correlates with *LIN28A*/*B* in cancer

Authors: Daniel J. Foster

Contributions: All experiments and analyses were performed by Daniel Foster.

Abstract

We analyzed *TRIM71* expression in databases of cancer patients as well as cancer cell lines finding that it is most highly expressed in hepatocellular carcinoma, testicular germ cell cancer, and acute myeloid leukemia. *TRIM71* expression correlates with *LIN28A*/*B* in hepatocellular carcinoma and testicular germ cell cancer indicating that the LIN28-TRIM71 genetic pathway may be intact and overexpressed in these cancer types. Despite seeing dependence on *TRIM71* in some liver cancer cell lines, patient survival is not correlated with *TRIM71* expression. Therefore, *TRIM71* may act as a bystander gene, downstream of LIN28 in cancer. Nonetheless, we pursued TRIM71 in liver cancer finding that depletion caused strong senescent phenotypes in HepG2 cells. Given the importance of TRIM71 in HepG2 cells, these cells were used for the CLIP studies to uncover important RNAs bound.

TRIM71 expression analysis and correlation with LIN28, CRD complex members, and RNA targets

TRIM71 TCGA expression and correlation with LIN28A/B

During development, TRIM71 is involved in the same genetic pathway as LIN28A/B and *let-7* (Figure 1.1B). *LIN28A/B* has been implicated as an oncogene in a vast number of cancers, while *let-7* has been implicated as a tumor suppressor. Due to the correlation between *LIN28A/B* and *TRIM71* during development we asked whether they were also correlated in cancer, and whether *TRIM71* can act as an oncogene.

We first looked at *TRIM71* expression across cancers by downloading patient data from The Cancer Genome Atlas (TCGA). We sorted by cancer type based on median *TRIM71* expression to find the cancers where it was most highly expressed. Testicular germ cell cancer (TGCC), acute myeloid leukemia (AML), and hepatocellular carcinoma (HCC) were the most highly *TRIM71* expressing cancers (Figure 3.1). *LIN28A/B* has been connected to each of these three cancers as an oncogenic driver when overexpressed (Thornton and Gregory, 2012). *LIN28A/B* expression leads to decreased levels of the *let*-7 microRNA leading to increased expression of *let*-7 targets like *MYC*, *RAS*, and *HMGA2* which are all oncogenes (Balzeau et al., 2017). Apart from its effect on *let-7*, LIN28 also binds to mRNAs like *IMP1* and *HMGA2* to positively regulate them. As these RNAs are

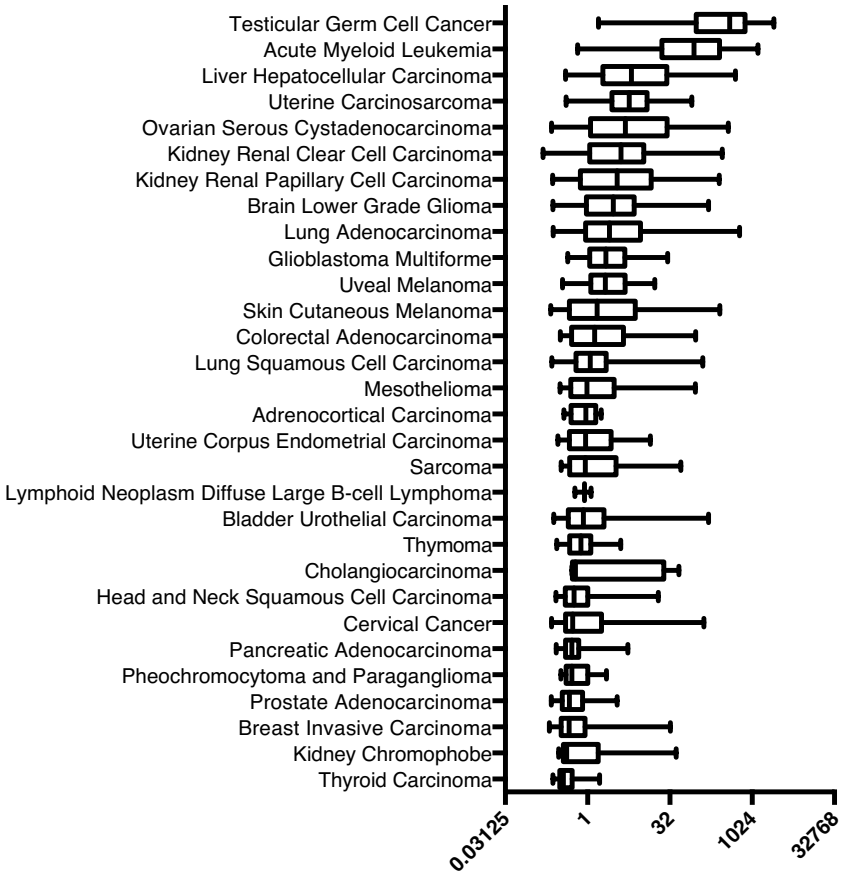


Figure 3.1. *TRIM71* expression by cancer type
TRIM71 expression (RSEM) is plotted for each cancer type with data available through The Cancer Genome Atlas (TCGA). Testicular germ cell, AML, and liver cancer are were most highly *TRIM71* expressing cancers.

both *let-7* and LIN28 targets, increased expression of *LIN28* ends up regulating them through dual mechanisms.

We next looked at the correlation between *TRIM71* and *LIN28A/B* in these cancers. *LIN28A* did not correlate with *TRIM71* expression in leukemia (data not shown). *LIN28B* was lowly expressed in leukemia patients based on TCGA data. Consequently, its expression did not correlate with *TRIM71*. In testicular cancer, *TRIM71* correlated strongly with both *LIN28A* and *LIN28B* (Figure 3.2A). In liver cancer patients *TRIM71* correlated with *LIN28A/B* as well (Figure 3.2B).

TRIM71 expression is higher in tumors compared to normal tissue

Although *TRIM71* appeared highly expressed in these cancers and correlated well with *LIN28A/B*, it's important to know whether the expression is related to the cancer itself or simply the tissue of origin. For example, *TRIM71* appears highly expressed at a basal level in adult testis based on Genotype-Tissue Expression (GTEx) data so its expression in testicular cancer may not be oncogenic. Unfortunately, TCGA does not have matched normal tissue for most patients. Using the UCSC Xena browser (Goldman et al., 2019), which compiles TCGA and GTEx datasets and renormalizes their expression for comparative purposes, we were able to look at *TRIM71* expression in the various cancers along with expression in normal tissue. Despite high *TRIM71* expression in testis, expression was still higher in testicular cancer. Likewise, expression in AML and HCC was higher than their respective tissues (Figure 3.2C).

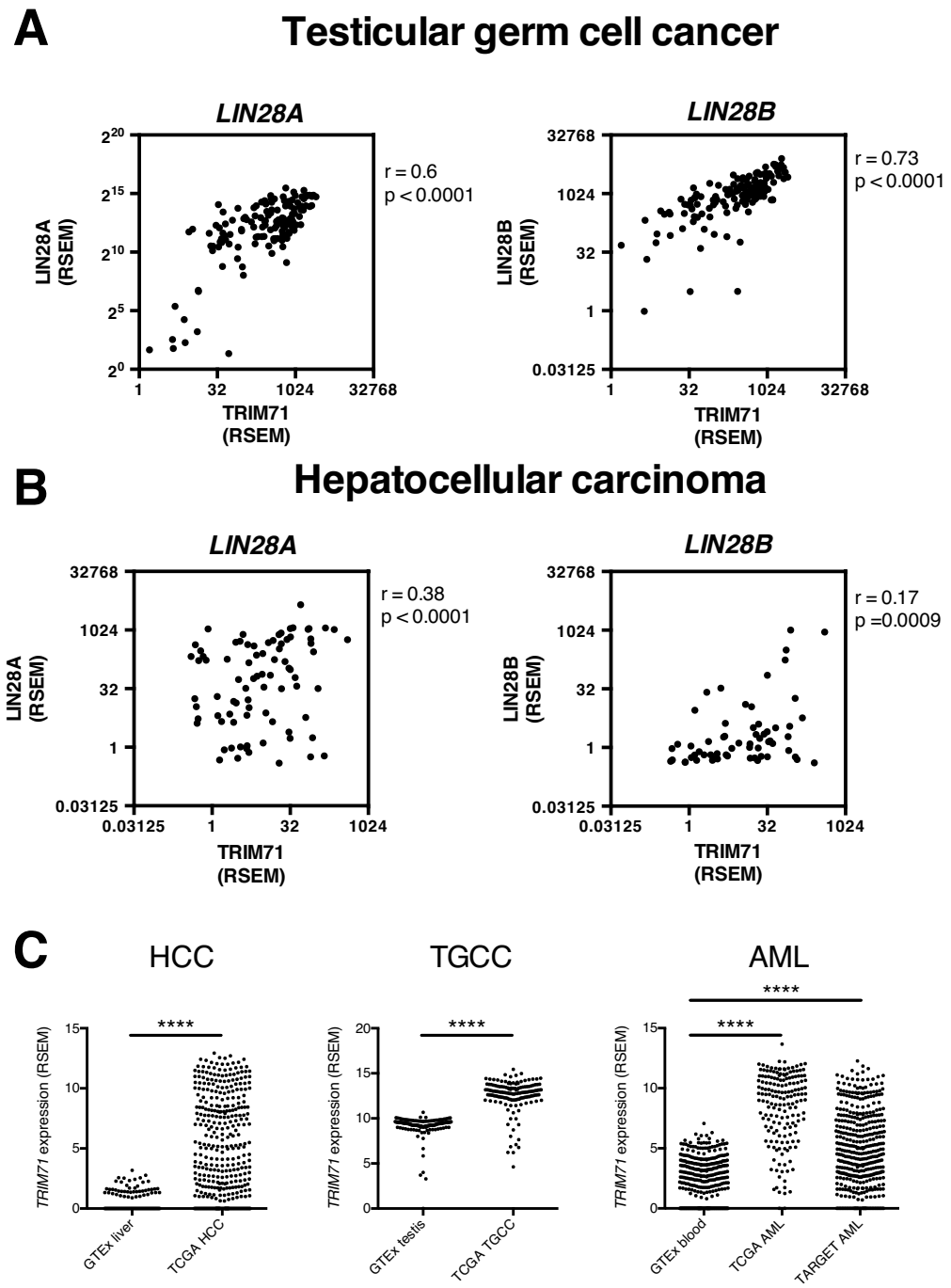


Figure 3.2. *TRIM71* expression positively correlates with *LIN28A/B* expression.

Scatterplots of *TRIM71* expression plotted against *LIN28A* or *LIN28B* expression in **(A)** testicular germ cell cancer or **(B)** hepatocellular carcinoma from TCGA. *TRIM71* significantly correlated with *LIN28A/B* in both cancer types. **(C)** *TRIM71* was expressed significantly higher in hepatocellular carcinoma, testicular germ cell cancer, and acute myeloid leukemia compared to GTEx expression data of comparative normal tissue. Mann-Whitney test **** represents p<0.0001.

TRIM71 expression correlates with CRD complex members and target RNAs.

Given than TRIM71 is able to participate in RNA binding with CRD complex members in HepG2 cells, we wanted to see if that interaction was possibly maintained in patients. We checked whether CRD members were expressed and whether they correlated with *TRIM71*. In liver cancer, *IMP1* and *DHX9* were expressed and did correlate with *TRIM71* (Figure 3.3A). In leukemia, *TRIM71* expression correlated with *DHX9* (Figure 3.3B), but *IMP1* expression was quite low and undetectable in most patients.

Next, we looked at expression of *MYC*, a target of the CRD complex that also plays an oncogenic role in liver cancer and leukemia. We found that the top quartile of *TRIM71* expressing patients had significantly higher *MYC* expression than the bottom 75% in both liver cancer and leukemia (Figure 3.3C). Although we've shown that TRIM71 binds the RNA and positively regulates *MYC*, that is not necessarily what leads to this correlation. In fact, *TRIM71* is a direct transcriptional target of MYC (Chen et al., 2013) which could also explain the correlation. Nonetheless, we wanted to follow-up on the possibility that TRIM71 could play an oncogenic role in these cancers.

TRIM71 in cancer cell lines

To study the role of TRIM71 in cancer we needed to find cell lines that expressed it (Klijn et al., 2015). Furthermore, we were interested in seeing whether cancer cell lines showed similar trends in *TRIM71* expression as patient samples. We downloaded an RNA-seq dataset from Klijn, et al., 2015, where 675 commonly used cancer cells lines were profiled. We categorized the cell lines based on their tissues of origin and found that

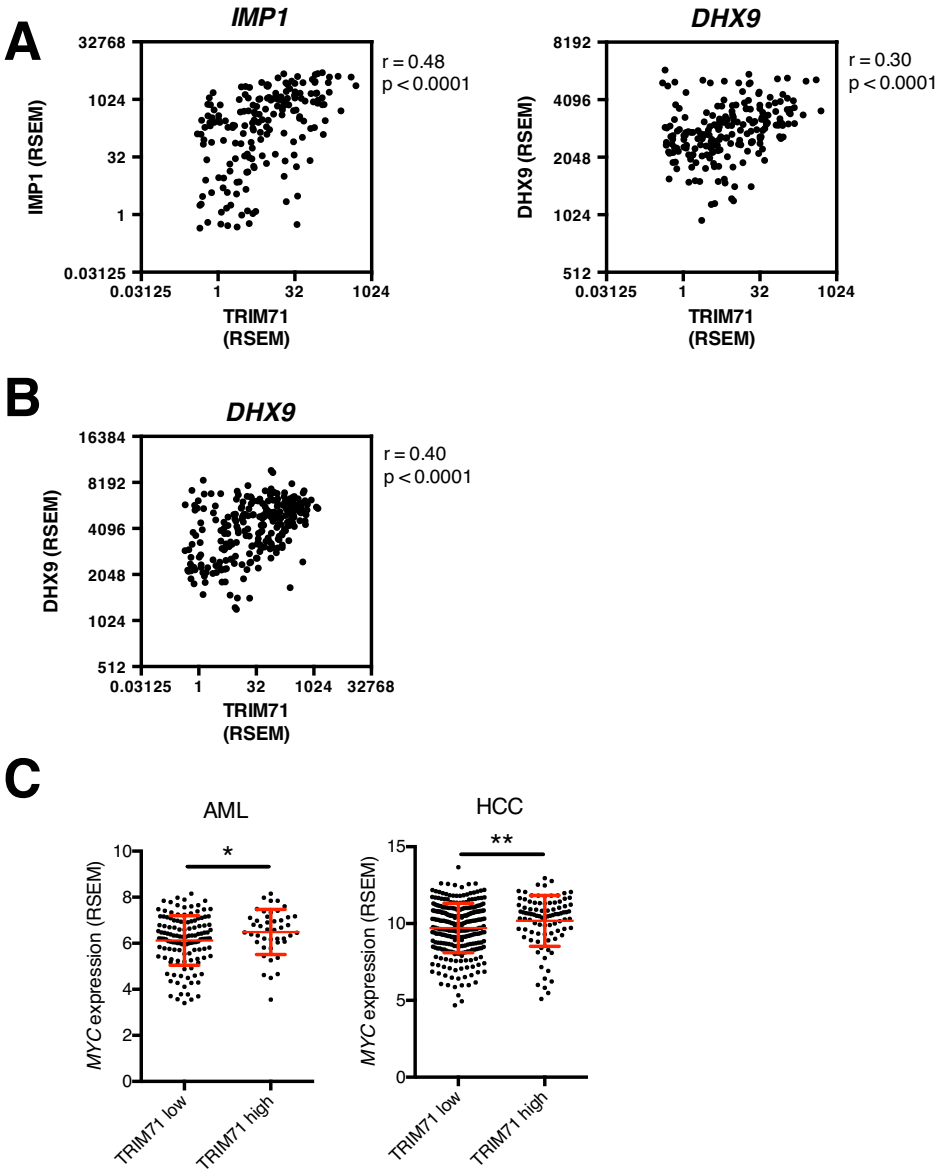


Figure 3.3. *TRIM71* expression correlates with *IMP1, DHX9,* and *MYC.*
(A) Scatterplots of *TRIM71* expression plotted against *IMP1* or *DHX9* expression in hepatocellular carcinoma. *TRIM71* was correlated with both *IMP1* and *DHX9* in HCC. **(B)** *TRIM71* and *DHX9* are correlated in AML. **(C)** Patients were grouped based on *TRIM71* expression with the upper quartile being compared against the lower 75%. *MYC* expression was compared and was significantly higher in the upper quartile of *TRIM71* expressing patients. Red lines represent the mean +/- SD. Mann-Whitney test, * and ** represent $p<0.05$ and $p<0.01$, respectively.

liver cancer cell lines were the most highly *TRIM71* expressing (Figure 3.4). Leukemia, which was categorized into lymphoid tissue, was third on the list of *TRIM71* expression. Sarcoma came up as second on the ranked list, however from the TCGA dataset sarcoma does not appear to express *TRIM71* highly (Figure 3.1).

Notably, most cancer cell lines had quite low *TRIM71* expression. In fact, the majority of cell lines had expression below an RPKM value of 1, which is a common cutoff to be considered as expressed. We next looked at *TRIM71* expression within specific cancer types we were interested in studying. We plotted expression for each hepatocellular carcinoma and acute myeloid leukemia cell line in addition to neuroblastoma for which *LIN28B* expression has been linked (Molenaar et al., 2012). We found that within HCC there was a subset of cells that expressed *TRIM71* more highly than the rest (Figure 3.4B). Within AML and neuroblastoma, there were also some cells that expressed *TRIM71* more highly.

As *LIN28A/B* has been implicated in many cancer types we wanted to know whether it was generally correlated with *TRIM71* across cancers. Looking at all 675 cancer cell lines, *TRIM71* expression does weakly correlate with *LIN28B* (Figure 3.4C). Next we looked at this correlation specifically within liver cancer cell lines and saw a strong positive correlation between the two genes (Figure 3.4D). This is similar to the correlation of *LIN28A/B* and *TRIM71* in HCC patient samples from TCGA (Figure 3.2B). Finally, looking at neuroblastoma cells lines, we again saw a strong correlation between *LIN28B* and *TRIM71*.

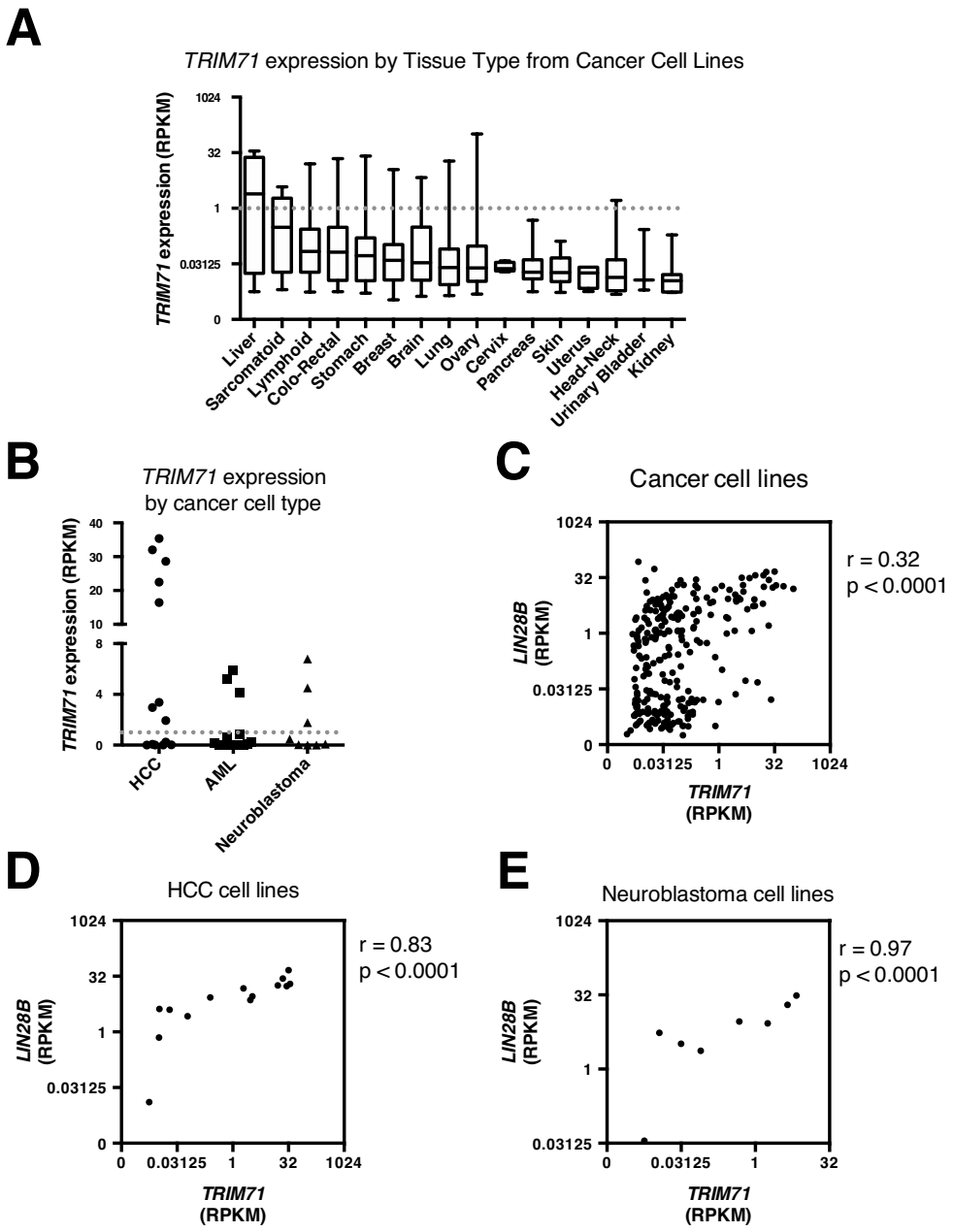


Figure 3.4. *TRIM71* expression and correlation with *LIN28B* in cancer cell lines. **(A)** *TRIM71* expression values were plotted for 675 cancer cell lines organized based on tissue type of origin. Liver cancer cells were the most highly *TRIM71* expressing cancer cell lines profiled. **(B)** Cells lines from cancers of interest were plotted based on *TRIM71* expression. Within hepatocellular carcinoma (HCC) there were two populations of cells, one that highly expressed *TRIM71* and one that did not. Within AML, the majority of AML cell lines did not express *TRIM71*. Among those that did, it was more lowly expressed than HCC. Within brain cancer cells, several neuroblastoma cell lines expressed *TRIM71*. **(C, D, E)** *TRIM71* correlated with *LIN28B* when looking at all 675 cell lines as a whole **(C)** and when focused on HCC **(D)** and neuroblastoma cell lines **(E)**.

The effect of TRIM71 expression on patient outcomes and growth of cell lines

TRIM71 correlation with survival

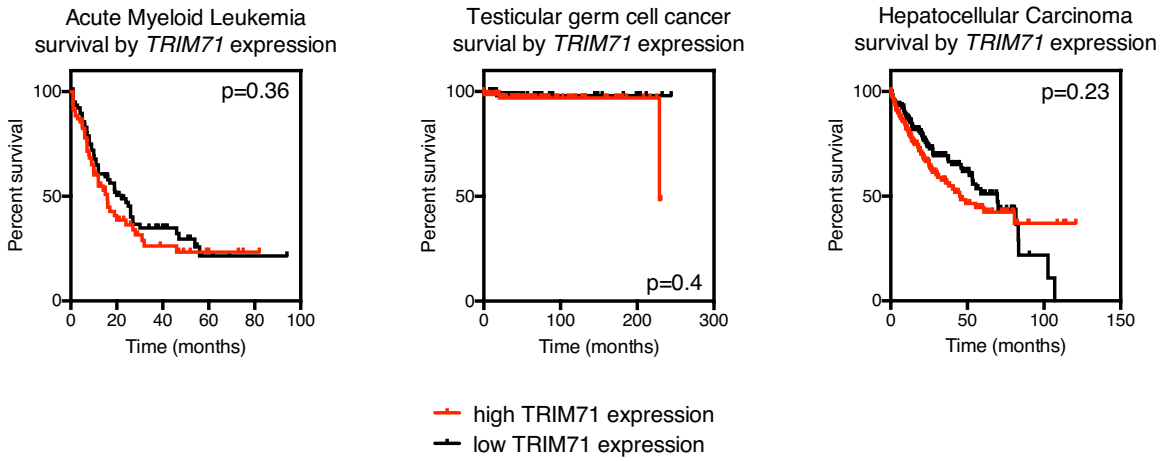
Patient survival data was downloaded from TCGA and sorted based on median *TRIM71* expression from each dataset. For AML and HCC, high *TRIM71* expression appeared to cause a small decrease in survival, though neither of these effects were significant (Figure 3.5A). For testicular cancer, survival rates were already quite high, and *TRIM71* expression did not appear to affect this. This indicates that TRIM71 could act as a bystander gene rather than an oncogenic driver.

We also wanted to look at the effect of *TRIM71* expression in other cancers besides those already mentioned. Using the KMplotter database which collects data from TCGA, survival was plotted based on *TRIM71* expression for many different cancer types. This analysis showed 6 cancers in which *TRIM71* expression correlated with decreased survival and 6 cancers in which it correlated with increased survival (Figure 3.5B). This analysis was performed based on most significant expression cutoffs rather than median expression from above and is prone to false positives. However, through this analysis, liver cancer did turn up as being associated with decreased survival based on *TRIM71* expression. Overall, the patient survival data based on *TRIM71* expression indicates that there may not be a strong effect in the cancers looked at.

TRIM71 cancer dependencies

Following our analysis of patients based on *TRIM71* expression, we next looked at cancer cell lines in the Broad Project Achilles database (Meyers et al., 2017; Tsherniak et al., 2017). The goal of project Achilles is to determine genetic dependencies and

A



B

TRIM71 correlates with decreased survival

	p-value
Sarcoma	0.00064
Kidney renal papillary cell	0.00160
Uterine corpus endometrial carcinoma	0.01300
Liver cancer	0.02800
Bladder carcinoma	0.03900
Esophogeal cancer	0.04700

TRIM71 correlates with increased survival

	p-value
Thymoma	0.00097
Thyroid carcinoma	0.00140
Cervical squamous cell carcinoma	0.00210
Kidney renal clear cell carcinoma	0.01400
Pancreatic ductal adenocarcinoma	0.01600
Head and neck squamous cell carcinoma	0.03000

Figure 3.5. Survival of high vs low *TRIM71* expressing patients.
(A) Patient survival data was downloaded from TCGA. Expression was stratified into low expressing and high expressing based on median *TRIM71* expression from each data set. Differences in survival based on *TRIM71* expression by median did not reach statistical significance for these three cancers. However, liver cancer did reach significance using different cutoffs to assign high vs low *TRIM71* expression. **(B)** *TRIM71* expression correlates with decreased survival in 6 cancer types and with increased survival in 6 cancer types. Data are from TCGA and is calculated using KMplotter based on most significant expression cutoffs.

essential genes by performing RNAi and CRISPR screens across the genome in ~2000 cancer cell lines. A dependency score (CERES) is assigned to each cell line for every gene, with a lower score meaning the cells are more dependent on the particular gene. Dependency scores are normalized to a scale where a CERES score of 0 signifies genes with no dependency, while -1 is the median score of all essential genes. Oncogenes such as *LIN28B* or *MYC* have scores below 0, whereas protective tumor suppressors such as TP53 have scores greater than 0.

We analyzed the project Achilles database for *TRIM71* and found that every cancer in the database had a mean CERES score below 0 (Figure 3.6, note reverse x-axis). Sorting cancers based on their mean score showed rhabdoid and liver cancers as the most dependent. With each point representing a different cell line, we saw that some cell lines were more dependent than others. As previously discussed, many cancer cell lines expressed low levels of *TRIM71* (Figure 3.4A). Therefore, we tested whether the CERES scores correlated with *TRIM71* expression. Perhaps unsurprisingly, higher *TRIM71* expression did correlate with stronger dependency on *TRIM71* (Figure 3.6B). Next, we looked at expression vs dependency for liver cancer, AML, and neuroblastoma (Figure 3.6C-E). We found a strong correlation in liver cancer and AML, though expression was lower in AML compared to liver, and the dependencies were more modest. *TRIM71* expression was not significantly correlated with dependencies in neuroblastoma (Figure 3.6E).

Finally, we looked at *LIN28B* dependency in these cancers as a positive control gene that had been previously validated. *LIN28B* expression correlated with dependency in liver (not shown) and neuroblastoma (Figure 3.6F). However, it did not correlate in AML (Figure 3.6G). Despite the Broad Achilles data, LIN28 has been shown to play a role in

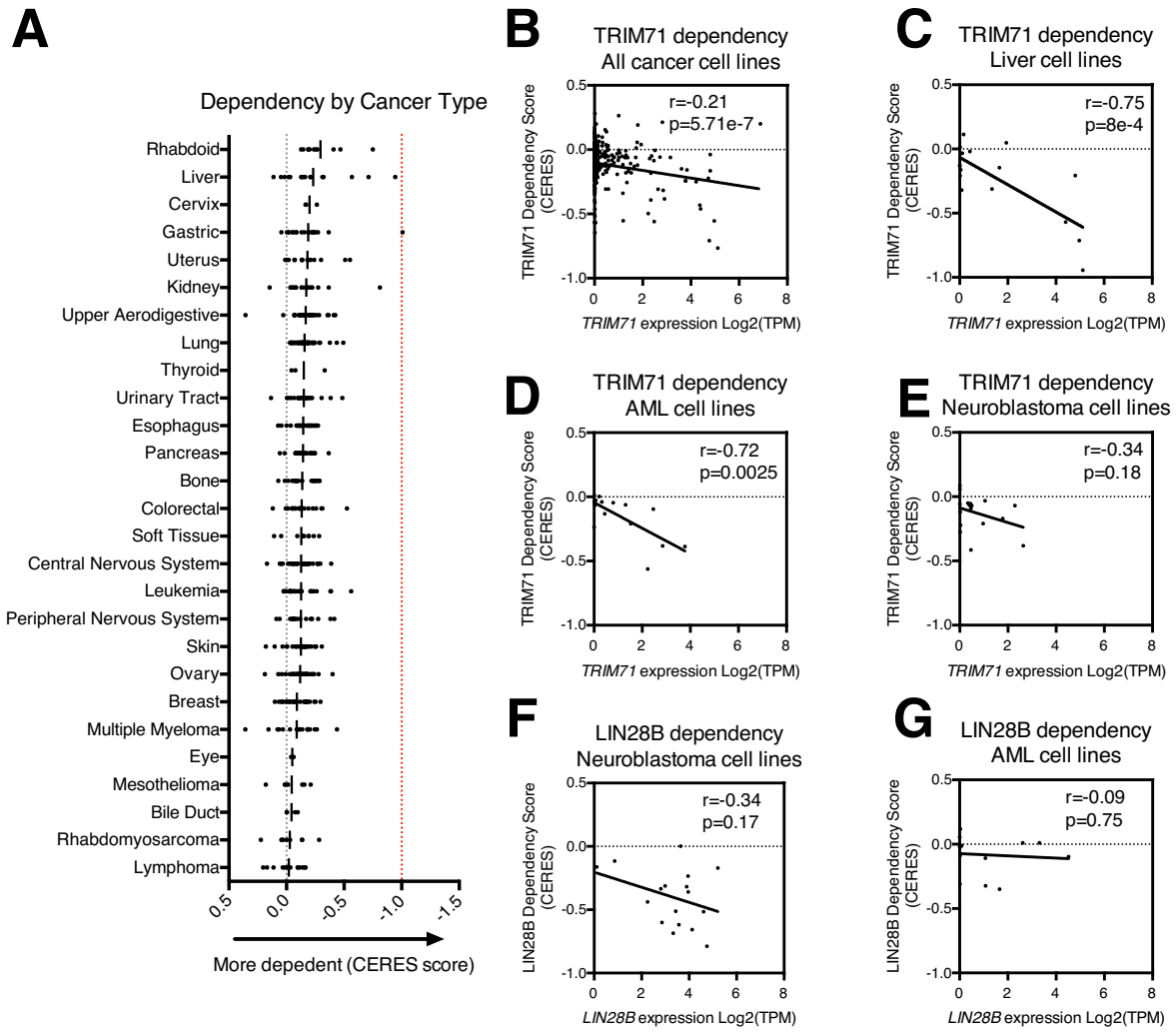


Figure 3.6. TRIM71 dependencies in cancer.
(A) Dependency score data from cancer cell lines was downloaded from Broad's Project Achilles. Cell lines were then categorized by cancer type and ranked based on the mean dependency score of all cell lines from a particular cancer type. A more negative score, towards the right, represents a stronger dependency on TRIM71. Cells with scores less than zero depend on expression of *TRIM71* suggesting it is an oncogene. A score of -1 is the median score of all pan-essential genes. Liver cancer cell lines were some of the most dependent cancer cell lines. **(B)** Among all cancer cell lines, dependency on TRIM71 was correlated with *TRIM71* expression. **(C, D)** Compared to all cell lines, in liver cancer and AML cell lines, TRIM71 expression was more strongly correlated with TRIM71 dependency. **(E)** In neuroblastoma cell lines, TRIM71 expression did not significantly correlate with dependency. **(F, G)** LIN28B dependency did not correlate with expression in neuroblastoma or AML. AML cell lines were generally not dependent on LIN28B while neuroblastoma lines were.

AML, so these results should be used only as a starting point requiring validation. Although *TRIM71* is expressed in AML it may not be related to the *LIN28A/B* developmental pathway since the two genes did not correlate in expression in TCGA patient data. In neuroblastoma, although *LIN28B* expression did not correlate with dependency, all cells were generally dependent on *LIN28B*.

Liver cancer was the only cancer type where *TRIM71* was expressed and correlated with *LIN28A/B* in both TCGA patients and cell lines. Further, *TRIM71* expression also correlated with CRD complex members and CRD mRNA target *MYC* in liver cancer. Lastly, liver cancer was one of the most dependent cancers based on Broad Achilles data. It was not predetermined that *LIN28A/B* should correlate with *TRIM71* as a requirement for *TRIM71* to be involved in cancer. However, correlations in the two provide a rationale for studying *TRIM71* and provide a potential explanation for how it becomes expressed. Additionally, in *LIN28B* overexpression mouse models of liver cancer, *IMP1* was shown to be a major downstream target of LIN28B and becomes expressed in response to LIN28B. Since we have now shown a connection between TRIM71 and IMP1, liver cancer became our highest priority going forward.

TRIM71 in liver cancer cell lines

We tested the effect of *TRIM71* depletion in several different liver cancer cell lines. We knocked down *TRIM71* with siRNAs and achieved ~95% knockdown at the RNA level in both HepG2 and Huh7 (Figure 3.7A). As shown previously (Figure 2.9) we also achieved good knockdown at the protein level. *TRIM71* depletion decreased cell proliferation as shown by MTT assay, a dye based colorimetric assay that indirectly measures the number of viable cells based on metabolic activity (Figure 3.7B). The effect on HepG2 cells was much greater than the effect on Huh7 cells in the MTT assay. Next

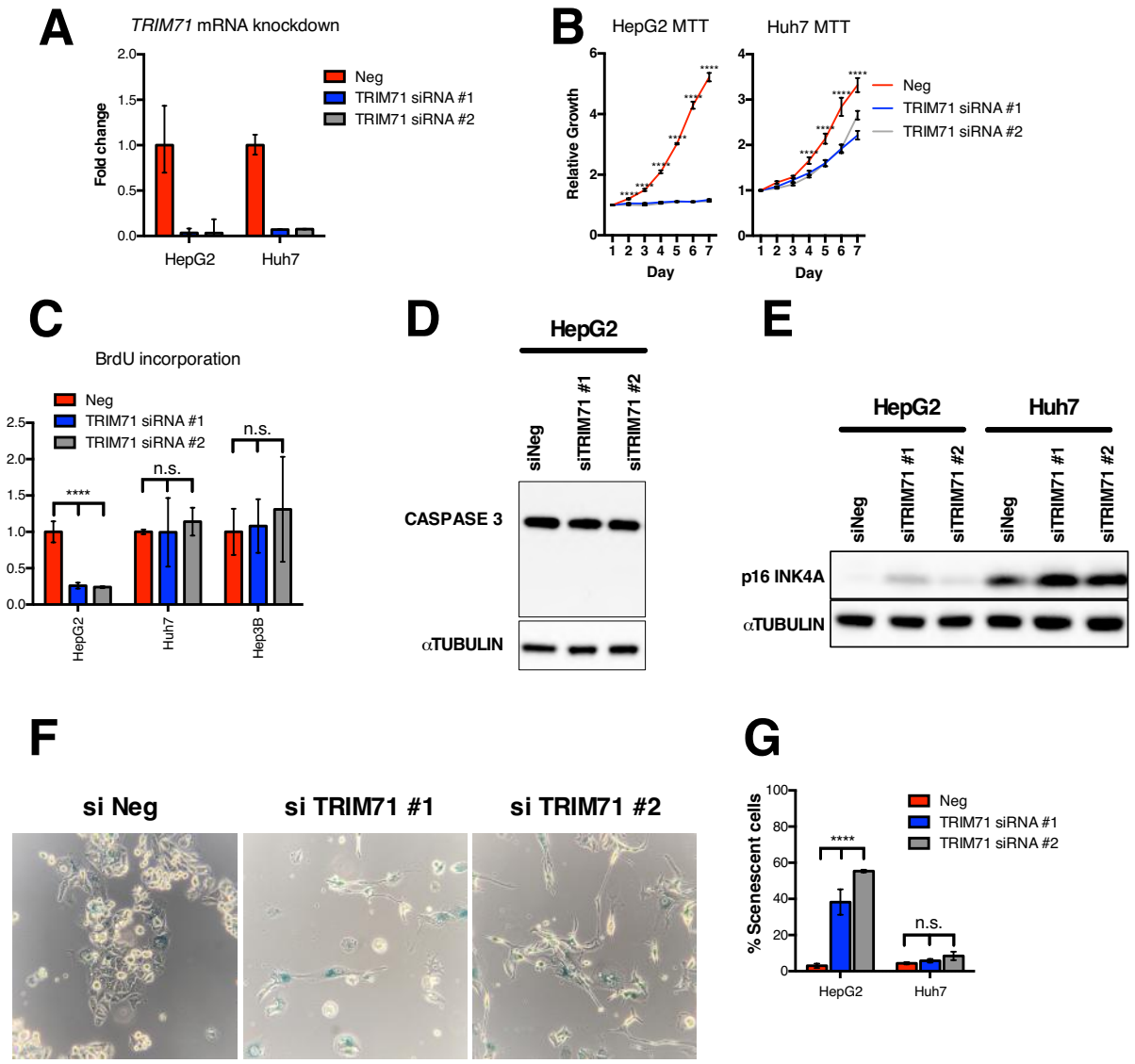


Figure 3.7. Knocking down *TRIM71* slows growth and induces senescence
(A) Knocking down *TRIM71* with siRNAs resulted in 90-95% knockdown at the mRNA level. **(B)** *TRIM71* knockdown decreased cell growth as shown by MTT assay. Error bars show S.D. of 3-5 biological replicates. 2-way ANOVA, $p<0.001$. **(C)** *TRIM71* knockdown decreased cell proliferation as shown by decreased BrdU incorporation into DNA in HepG2 cells but not Huh7 or Hep3B. One-way ANOVA, $p<0.001$. **(D)** *TRIM71* knockdown did not increase levels of cleaved Caspase-3 in HepG2 cells. **(E)** TRIM71 knockdown increased levels of senescence marker p16(INK4A) in HepG2 and Huh7 cells. **(F)** In HepG2 cells, *TRIM71* knockdown caused an increase in senescence associated β-galactosidase staining. These cells also exhibited an elongated morphological phenotype. **(G)** Quantification of part E showing increased SA β-gal staining in HepG2 cells. Huh7 cells did not show signs of β-gal staining. Error bars show S.D. of 3 biological replicates. 500-600 cells were counted per condition for each replicate. One-way ANOVA, $p<0.001$.

we measured BrdU incorporation, a DNA analog used to measure actively proliferating cells. *TRIM71* depletion decreased BrdU incorporation in HepG2 but not Huh7 or Hep3B cells.

We looked at both apoptosis and senescence to determine why *TRIM71* depletion decreased proliferation. First looking at apoptosis, we did not see any cleavage of Caspase-3 in HepG2 cells (Figure 3.7D). Next, we measured senescence by western blot for levels of p16(INK4A), a tumor suppressor involved in initiating senescence by pausing cells in the G1 phase of the cell cycle (Coppé et al., 2010; Ohtani et al., 2004). Both HepG2 and Huh7 showed a small increase in p16(INK4A) protein levels with *TRIM71* depletion indicating activation of senescence (Figure 3.7E). We also tested for Senescence-associated β-galactosidase (SA-β-gal) activity, another marker of cellular senescence. HepG2 cells showed strong SA-β-gal activity but Huh7 cells did not (Figure 3.7F, G). Additionally, *TRIM71* depletion also caused a strong change in cell morphology, with cells becoming elongated. These elongated cells were mostly senescent (Figure 3.7F).

To understand why there was a difference in phenotypes between HepG2, Huh7, and Hep3B we looked at genotypic differences. Liver cancer can be broadly characterized by subtype into mutually exclusive mutations in either *TP53* or *CTNNB1*, the gene encoding β-catenin (Guichard et al., 2012; Schulze et al., 2015). HepG2 cells have a *CTNNB1* mutation and are wildtype for p53. Conversely, Huh7 and Hep3B have a wild-type Wnt pathway and mutant p53. Considering that p53 plays an important role in the cell cycle as a tumor suppressor, it has also been described as being crucial for induction

and maintenance of cellular senescence (Qian and Senescence, 2013; Rufini et al., 2013). Therefore, we would not expect strong induction of cellular senescence in Huh7 or Hep3B.

Some studies have characterized HepG2 cells as a hepatoblastoma as opposed to hepatocellular carcinoma because it was derived from a 15 year old patient (Mihara et al., 1997). Hepatoblastoma is a childhood liver cancer that typically develops between 1-2 years of age. Similar to HCC, it is also commonly driven by mutations in *CTNNB1* and the Wnt pathway (Aguiar and Carneiro, 2017; Koch et al., 1999). Not many hepatoblastoma cells lines exist, however, we were able to study Huh6 cells which, similar to HepG2, are p53 wild-type and driven by mutations in *CTNNB1*. In Huh6 cells we were able to deplete *TRIM71* and look at prospective RNA targets (Figure 2.9) however we were not able to perform any cancer cell phenotypic assays as the cells were exceptionally difficult to grow in culture. Subjectively, knockdown of *TRIM71* did not appear to cause a phenotype, however, cells often require splitting after siRNA depletion before seeing any phenotypes, and this was not possible with these cells.

TRIM71 depletion causes a small effect on the Wnt pathway

Given the possible importance of TRIM71 in *CTNNB1* mutant cells, we looked at whether *TRIM71* expression was driven by β-catenin. Depletion of *CTNNB1*, the gene encoding β-catenin, by siRNAs did in fact decrease *TRIM71* expression (Figure 3.8A). It is possible that *TRIM71* could be a direct β-catenin target. Otherwise, the effect could be mediated by MYC, a known β-catenin target that transcriptionally regulates *TRIM71*. We have confirmed that MYC is able to regulate *TRIM71* in HepG2 cells (Figure 3.8A). Next,

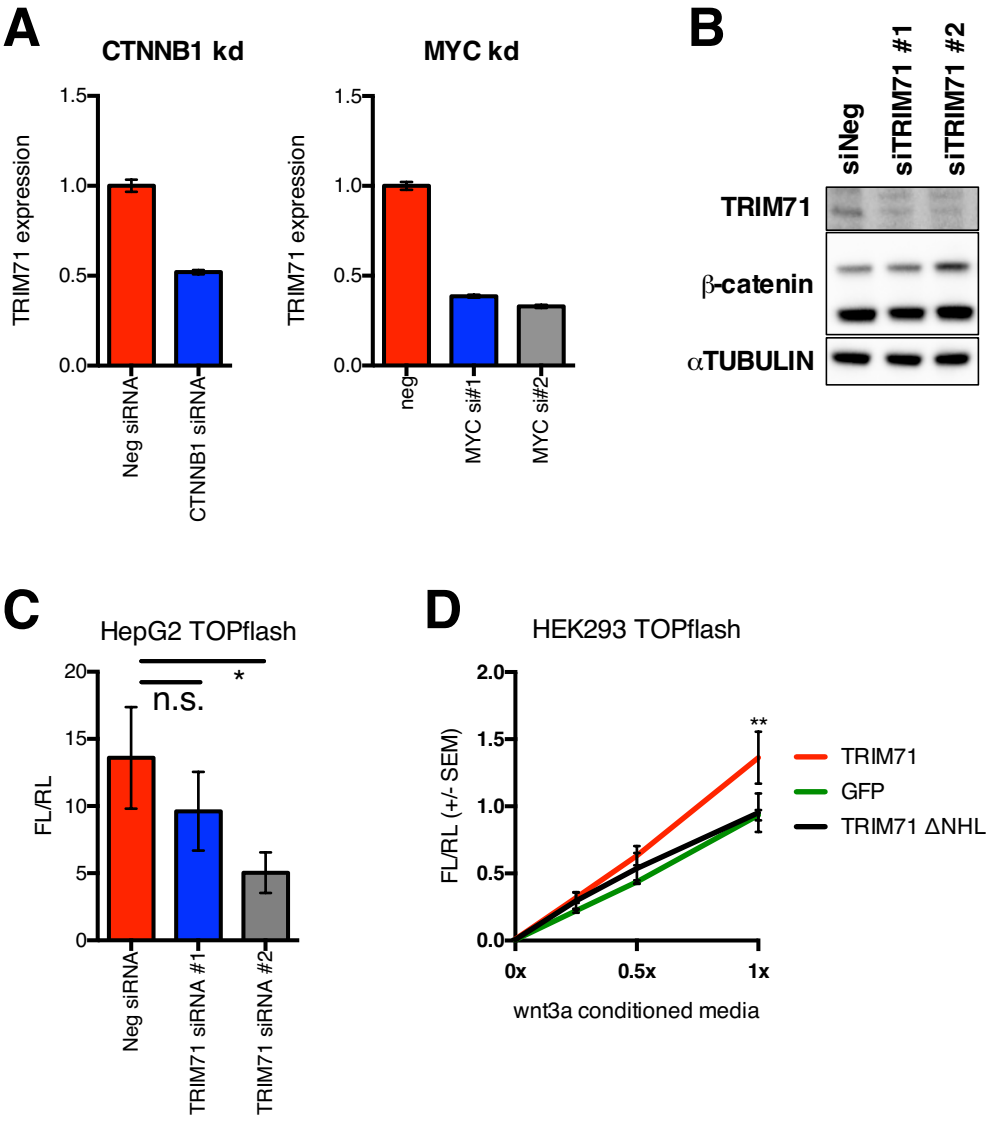


Figure 3.8. TRIM71 has a small effect on β-catenin activity.
(A) Depletion of CTNNB1 and MYC decreased levels of *TRIM71* expression. **(B)** TRIM71 depletion did not affect β-catenin levels. **(C)** TRIM71 depletion may have a small effect on TOPflash activity in HepG2. Only one TRIM71 siRNA showed an effect. **(D)** In HEK293 cells, overexpression of full-length TRIM71 but not an NHL deletion mutant or GFP potentiated TOPflash activity when cells were stimulated with wnt3a conditioned media.

we looked at whether TRIM71 affected β-catenin. However, *TRIM71* depletion did not appear to impact it at the protein level (Figure 3.8B). HepG2 cells have a deletion in the N-terminal region of β-catenin which leads to loss of regulation and constitutive activity. This explains the two β-catenin bands apparent in the western.

Active β-catenin binds to TCF/LEF transcription factors that mediate transcriptional activation caused by Wnt signaling. Since we've already shown that TRIM71 is able to bind and promote *TCF7* and *TCF7L2* (Figures 2.8, 2.9), we wondered whether TRIM71 could regulate TCF associated transcription. To test this, we transfected cells with a TOPflash luciferase reporter construct that contains 8x TCF/LEF binding sites in the promoter region of firefly luciferase. Since HepG2 cells have constitutively active β-catenin they have high reporter activity. We measured luciferase activity in *TRIM71* depleted cells and noticed a possible small decrease in activity. However, only one siRNA caused a significant decrease in luciferase activity while the other did not.

Next, we tested whether TRIM71 could facilitate TOPflash activity in HEK293 cells which have low baseline activity. To activate the TOPflash reporter, cells were grown in the presence of Wnt3a conditioned media that contains recombinant Wnt3a protein produced from cultured mouse fibroblasts. The media was added in dilutions to achieve a dosage effect. In the presence of overexpressed *TRIM71* but not *GFP*, TOPflash activity was increased at the highest Wnt3a dosage (Figure 3.8D). We also overexpressed a *ΔNHL TRIM71* mutant to test whether the effect was dependent on TRIM71 RNA binding. The ΔNHL did not facilitate activity indicating that the effect is likely RNA-binding dependent.

It appears that TRIM71 may play a small role in facilitating β-catenin/TCF transcription. Although TRIM71 does not directly affect the levels of β-catenin, it can directly bind to and promote *TCF7* and *TCF7L2*. This could explain why *TRIM71* depletion cause a small change in TOPflash activity. However, this effect was small and it may not explain the growth defect seen with *TRIM71* depletion. Therefore, we followed up on other TRIM71 targets.

TRIM71 in neuroblastoma cell lines

Given the correlation between *TRIM71* and *LIN28B* in neuroblastoma cells lines and the oncogenic role of *LIN28B* in neuroblastoma, we tested whether *TRIM71* also had an oncogenic role. *LIN28B* expression has been linked with *MYCN* (Molenaar et al., 2012) and we confirm that we see higher *LIN28B* in *MYCN* amplified cells (Figure 3.9A). Likewise, we see higher *TRIM71* expression in these same cells. We depleted *TRIM71* with shRNAs in these cells but only achieved modest knockdown of 20-60%, depending on the shRNA and cell line (Figure 3.9 B). Next, we counted cells to measure whether *TRIM71* depletion caused a growth defect but we did not see any change in growth (Figure 3.9C). Even focusing on the shRNAs with better knockdown, there was no effect. Therefore, although *TRIM71* and *LIN28B* are correlated in neuroblastoma, TRIM71 does not affect the growth of the cells. Thus, is seems that TRIM71 acts as a bystander in neuroblastoma as opposed to a driver. However, it is possible that better *TRIM71* depletion could have mediated an effect.

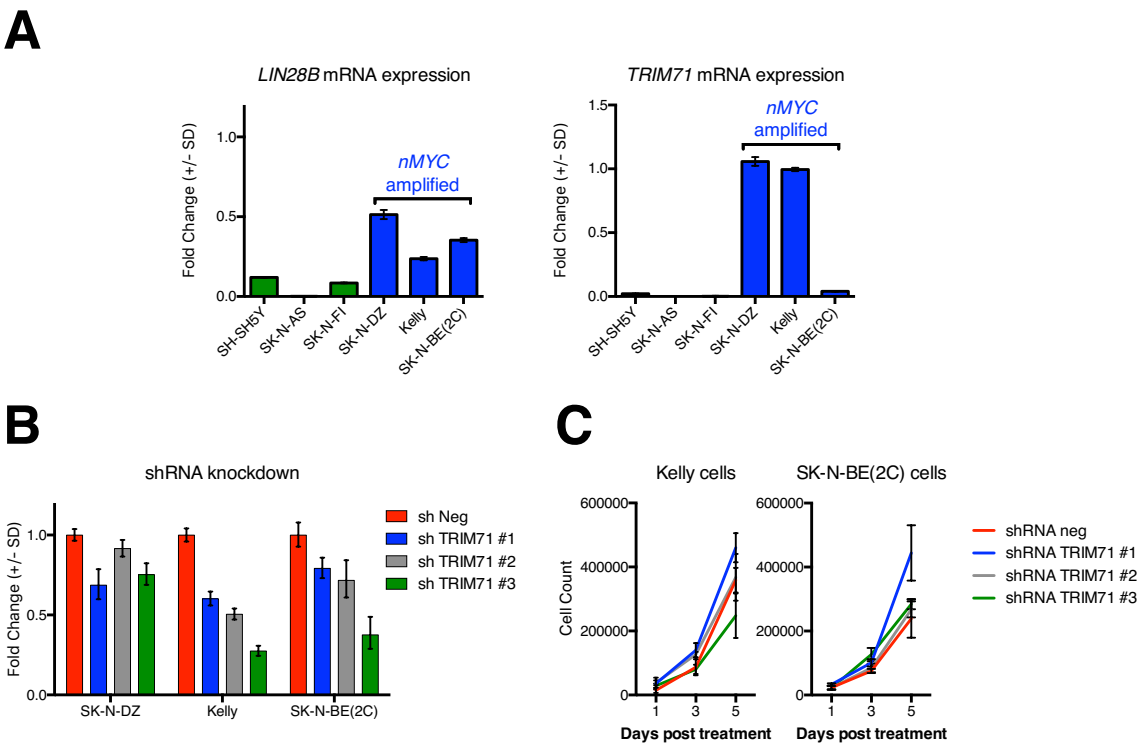


Figure 3.9. TRIM71 expression and knockdown in neuroblastoma cell lines.
(A) Both *TRIM71* and *LIN28B* were more highly expressed in *nMYC* amplified neuroblastoma cell lines compared to non-amplified cell lines. Fold change values are normalized to expression in HEK293 cells. **(B)** *TRIM71* shRNA knockdown in *nMYC* amplified cells gave between 20-60% knockdown. **(C)** *TRIM71* knockdown did not significantly slow growth in Kelly or SK-N-BE(2C) cells.

Rescuing *TRIM71* phenotypes in liver cancer cell lines

CDKN1A is one of the most reproducible TRIM71 mRNA targets. Because TRIM71 inhibits *CDKN1A*, *TRIM71* depletion causes an increase in p21 levels. Although *TRIM71* knockdown did not cause a very large increase in p21 at the protein level (Figure 2.9A), we tested whether co-knockdown of *TRIM71* and *CDKN1A* could rescue the growth phenotype on HepG2 cells. Despite good *TRIM71* depletion and moderate *CDKN1A* depletion (Figure 3.10A), the growth phenotype was not rescued as measure by MTT assay (Figure 3.10B).

We also tested whether we could rescue phenotypes by overexpressing mRNAs promoted by TRIM71. We overexpressed *GFP*, *LIN28B*, *MYC*, or *MDM2*, and tested for senescence. *MYC* overexpression rescued the morphological change apparent with *TRIM71* depletion (Figure 3.10C). *MYC* also partially rescued senescence caused by one *TRIM71* siRNA but not the other (Figure 3.10D). Neither *LIN28B* nor *MDM2* overexpression rescued the senescence or morphology. It is possible that the phenotypes associated with *TRIM71* depletion were MYC dependent. However, we cannot rule out the possibility that *MYC* overexpression generally facilitated cell growth and resistance to senescence, independent of *TRIM71* levels.

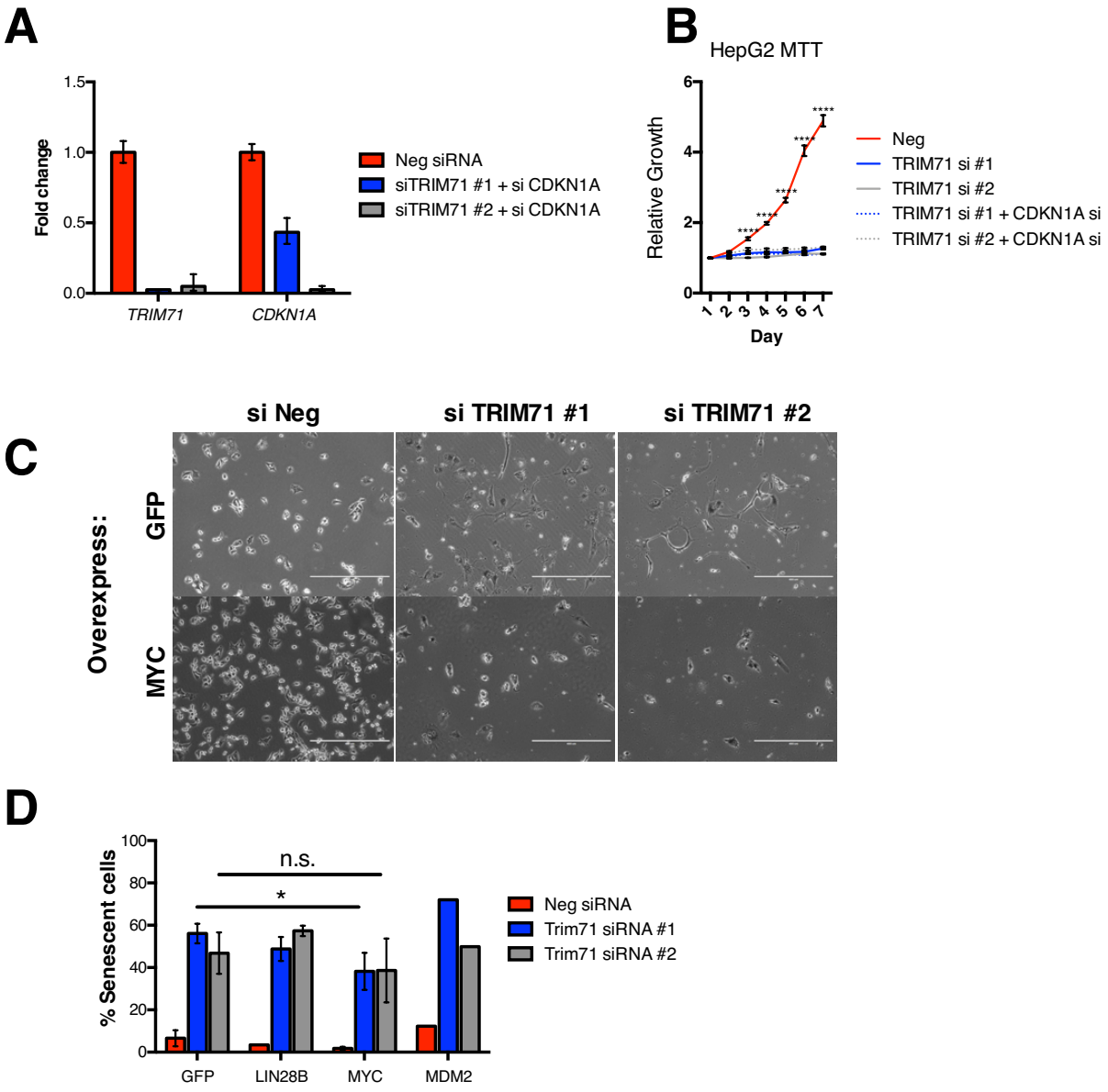


Figure 3.10. Rescuing *TRIM71* knockdown phenotypes.
(A, B) Co-knockdown of *TRIM71* and *CDKN1A* did not suppress the *TRIM71* knockdown phenotype. **(C)** Overexpression of *MYC* but not *GFP* suppressed the morphological phenotypes associated with *TRIM71* knockdown. However, *MYC* overexpression did rescue the growth phenotype as there are fewer cells in si TRIM71 vs si Neg. **(D)** Overexpression of *LIN28B* or *MDM2* did not rescue the senescence phenotype of *TRIM71* knockdown. Overexpression of MYC rescued for TRIM71 siRNA #1 but not #2. Error bars represent SD. Two-way ANOVA $p<0.05$. MDM2 samples only represent one biological replicate.

Summary

We started this analysis of TRIM71 in cancer by asking whether the LIN28/*let-7* developmental pathway was maintained in cancers where *LIN28A/B* was important as an oncogenic driver. Based on correlations in expression between *LIN28A/B* and *TRIM71* in both TCGA patient samples as well as cancer cell lines, the developmental pathway does seem to be maintained. Next it was important to determine whether *TRIM71* expression was associated with poor patient survival or decreased growth of cancer cell lines. The overall survival of patients separated based on high versus low *TRIM71* expression levels was not significantly different. This suggests that TRIM71 may not be an oncogenic driver. Despite this, we pursued studying TRIM71 in the context of liver cancer cell lines due to their strong dependency in the Broad Achilles dataset. We found strong growth defects when *TRIM71* was depleted in HepG2 and Huh7 cells. However the growth defects were stronger in HepG2 cells which have an overactivated Wnt pathway due to a *CTNNB1* mutation.

Future studies of TRIM71 in cancer should dive deeper into the potential connections between TRIM71 and the Wnt pathway. Interestingly, the Wnt pathway is important for development of the liver. Next steps will involve overexpressing TRIM71 in the liver to determine whether it is sufficient to transform cells.

Chapter 4 – Discussion

Model for understanding TRIM71 as a member of the RNA-binding CRD complex

In chapter 2 we showed that TRIM71 was capable of co-precipitating with CRD complex members IMP1 and DHX9 (Figure 4.1A). These are both RNA binding proteins with their own RNA binding domains. We determined that the coiled-coil domain of TRIM71 was responsible for the interaction with DHX9 but we could not narrow the IMP1 interaction to a particular domain. We did however show that the RNA-binding NHL domain of TRIM71 did not interact with either protein on its own. CLIP-sequencing showed that 30% of the TRIM71 bound RNAs were also previously found to be bound by IMP1.

We hypothesized that if TRIM71 was a member of this complex that it would likely impact the RNAs bound (Figure 4.1B). To test this, we depleted *IMP1* and performed CLIP-qPCR for some of the RNAs discovered in the CLIP-seq. Some TRIM71 targets were enriched while others did not change. Interestingly, for a non-target RNA, *ACTIN*, we saw increased binding by TRIM71 when *IMP1* was depleted. We hypothesize that when CRD complex members are depleted, the specificity of RNA binding by TRIM71 goes down. However, we only tested this hypothesis with *IMP1* depletion and also only tested select RNAs. We did not test whether *IMP1* depletion dissociates the complex and whether DHX9 can still bind to TRIM71 (Figure 4.1B) but these would be informative future experiments.

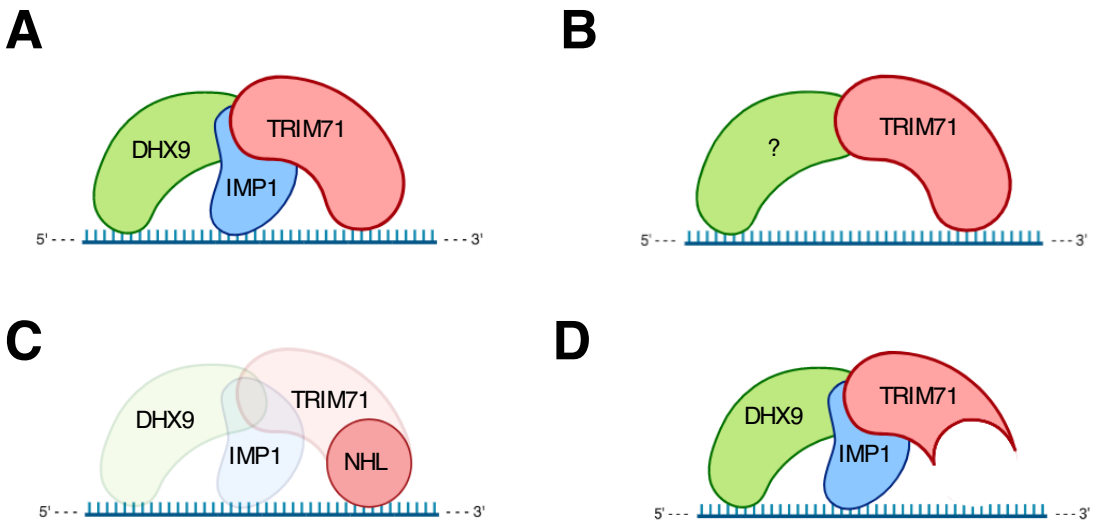


Figure 4.1. CRD complex members provide specificity of binding for TRIM71.
(A) Schematic of RNA bound by TRIM71 and the CRD complex. Note there are multiple proteins in the complex with their own RNA-binding motifs. **(B)** Depletion of complex member IMP1 decreases the specificity of TRIM71 binding. **(C)** The NHL domain on its own is unable to bind the CRD complex members, represented as transparent. In this scenario the NHL domain is binding based only on RNA motifs it recognizes. The lack of protein binding partners causes the NHL domain to bind with less specificity. **(D)** Deletion of the TRIM71 NHL domain decreases the ability of TRIM71 to bind target RNA, but it is able to bind RNAs of the CRD complex moreso than non-CRD targets when the NHL domain is deleted.

We next tested this hypothesis by performing CLIP-sequencing with the NHL domain of TRIM71. As previously shown, the NHL domain cannot precipitate with CRD complex members (Figure 4.1C). We hypothesized that if the CRD complex was specifying TRIM71 RNA binding that the NHL would show promiscuous binding. We did in fact show that the NHL binds to more RNAs than full-length TRIM71. Further, the full-length TRIM71 target RNAs were bound less strongly in the NHL-only CLIP. This suggests that TRIM71 relies on binding partners to help specify target RNAs and to bind them with high affinity.

Lastly, we tested RNA binding of a ΔNHL TRIM71 mutant. As expected, deletion of the RNA binding domain decreased the binding to target RNAs. However, binding to known IMP1 targets was not affected to the same extent. We showed that the ΔNHL TRIM71 mutant was still able to bind to CRD complex members. We suspect that this allowed TRIM71 to continue binding RNA targets of the CRD complex (Figure 4.1D).

Overall, our results from chapter 2 indicate that the RNA binding activity of TRIM71 should not be looked at in isolation. Rather, we have shown that protein binding partners impact RNA targets in a meaningful way.

TRIM71 is capable of positive or negative regulation of target RNAs

Previously it had been shown that TRIM71 was capable of inhibiting the mRNAs *CDKN1A* and *EGR1*. It was hypothesized that TRIM71 inhibits RNAs because tethering to luciferase reporters repressed reporter activity. We have confirmed that TRIM71 is capable of repressing these RNAs and others. However, given that the number of RNAs

depleted with *Trim71* knockout in mES cells is greater than those enriched, we hypothesized that TRIM71 may be able to promote RNAs.

We performed CLIP-sequencing and compared the bound RNA with those changed in the *Trim71* knockout mES cells and found more bound RNAs were depleted with *Trim71* knockout than enriched. Among the RNAs that were depleted in the *Trim71* knockout mES cells and bound by TRIM71 were oncogenes such as *LIN28B*, *MDM2*, *TCF7* and *TCT7L2*. We confirmed by CLIP-qPCR that TRIM71 was binding these RNAs and also promoting their expression at the protein level. We also examined the canonical CRD complex target *MYC* and saw enrichment in the CLIP-qPCR. Further analysis of the *MYC* RNA showed enrichment in the CRD element, a regulatory element within *MYC* where the CRD complex binds. Among these new RNA targets we tested for steady state RNA levels and found decreases in *LIN28B* and *MDM2* but not *MYC*. This indicates that TRIM71 could be affecting the stability of some RNAs and the translation of others.

Given that *LIN28B* and *MYC* are *let-7* targets we tested whether *TRIM71* depletion could be affecting *let-7* levels or activity. We did not see changes in *let-7a/b/c* levels. We also tested for *let-7* activity to see whether the *TRIM71* mRNA could be acting as a sponge for *let-7* which could account for increases in *let-7* targets. However, we did not see any changes in *let-7* luciferase reporter activity with *TRIM71* depletion. This suggests that the effects we observed with *TRIM71* depletion were due to a direct regulatory effect on the RNAs as opposed to an indirect effect through *let-7*.

CLIP-sequencing suggested that there were other TRIM71 target mRNAs that were likely inhibited. For example, we found that TRIM71 was binding to all members of the muscleblind like (*MBNL*) family and the nuclear factor one (*NFI*) family and inhibiting expression at the protein level of MBNL1/3. The *MBNL* family of splicing factors and the

NFI family of transcription factors are both important during differentiation and development. They could account for the neural phenotypes associated with the loss of *TRIM71* in mice and patients. *MBNL2* is highly expressed in the brain and is associated with neural development (Charizanis et al., 2012; Sznajder et al., 2016). The *MBNL* family promotes adult specific splicing patterns of a vast number of genes. Loss of *TRIM71* would presumably cause an increase in expression in the *MBNL* family and lead to premature adult splicing patterns which could negatively impact the timing of developmental patterns.

NFIA/B/X are all highly expressed in the developing brain, specifically in areas of active neuronal cell division and migration within the ventricular zone and cortical plate (Plachez et al., 2008). The *NFI* family regulates neuronal development and migration. These three family members have severe effects when deleted in mice or patients (Heng et al., 2012; Lu et al., 2007). Similar to the *MBNL* family, loss of *TRIM71* would presumably lead to premature expression of the *NFI* members would could explain the premature differentiation of neurons associated with *TRIM71* loss.

TRIM71 NHL mutation in patients

TRIM71 mutations have been associated with congenital hydrocephalus, a disease in which cerebrospinal fluid (CSF) builds up in the ventricles (Furey et al., 2018). This disease can arise from neural tube defects during development. Mice with mutations in *Trim71* have severe neural tube defects (Chen et al., 2012; Schulman et al., 2008). In the congenital hydrocephalus patients that have *TRIM71* mutations, only one allele is mutated. Therefore, this could represent either a haploinsufficiency, in which the one wild-type copy is not enough to produce the normal phenotype, or it could represent a gain of

function mutation. Notably, all the mutations occurred within the RNA binding domain of TRIM71 suggesting the RNA binding activity was important for the phenotypes.

First, we tested whether these particular mutations affected RNA binding. With a CLIP-sequencing experiment we showed that the NHL mutants clustered closely with the ΔNHL mutant indicating that these NHL point mutants are likely binding fewer RNAs or not binding RNAs as strongly. This suggested a loss-of-function effect by these mutations. When we focused on binding to RNA targets discovered in the wild-type TRIM71 CLIP-seq, we were able to show that the mutants were still able to bind most RNA targets. However, these RNAs were significantly less enriched in the NHL mutants CLIP-seq experiments. Although we did not directly test RNA affinity, decreased enrichment of target RNAs suggest decreased binding to targets. This could result in a haploinsufficiency effect by these *TRIM71* mutations. We think the *MBNL* family of splicing factors could be interesting to follow up on in these mutants to see whether loss of *MBNL* regulation could cause the neural phenotypes observed. Of the three family members, *MBNL2* was the least enriched in the *TRIM71* point mutant CLIPs. Interestingly, *MBNL2* is also associated with the strongest neural defects in mice, as mentioned above.

The mutations seen in patients are arginine to histidine point mutations. Arginine has a strong positive charge and is capable of interacting with the negatively charged backbone of RNA. Histidine is also positively charged, though it's not quite as polar as arginine. We've shown that these mutants are still able to bind TRIM71 target RNAs, although the interaction appears weaker. So, we wondered whether they could represent a gain of function effect and bind new RNA targets. A gain-of-function effect could account

for less binding to wild-type TRIM71 target RNAs if the new target RNAs are sequestering the protein. We found that there were 307 RNAs bound by both TRIM71 point mutants that were not bound by wild-type TRIM71. Regulation of these RNAs by TRIM71 could represent a gain of function effect. Therefore, these data support both loss-of-function and gain-of-function mechanisms of the NHL point mutations. Follow-up studies in neural progenitor cells could give insight into which mechanism predominates. Electromobility shift assays (EMSA) with competing probes can be used to measure the dissociation constants of protein-RNA interactions which would give insight into the affinity of TRIM71 for target RNAs.

TRIM71 in cancer

We performed an analysis of *TRIM71* with datasets from The Cancer Genome Atlas (TCGA) finding that it is highly expressed in Acute Myeloid Leukemia (AML), Testicular Germ Cell Cancer (TGCC), and Hepatocellular Carcinoma (HCC). In all three of these cancers *TRIM71* expression was higher than comparative normal tissue. However, *TRIM71* expression did not significantly correlate with survival in these cancers. We next looked for correlations between *TRIM71* and *LIN28A/B* expecting that they may be correlated given their co-expression patterns seen in development. We saw correlation in HCC from both the TCGA dataset as well as in cancer cell lines. From the Broad Achilles dataset which analyzes genetic dependencies in cancer, liver appeared to be one of the most dependent on *TRIM71*. However, patient survival was not correlated with *TRIM71* expression indicating that it may act as a bystander gene.

We followed up on *TRIM71* in liver cancer by depleting it in several cell lines. We observed a strong growth defect and senescence in HepG2 cells and a weaker growth

defect in Huh7 cells. This difference in the magnitude of the phenotypes could be explained by the mutations driving growth in these particular cell lines. HepG2 cells are driven by mutations in *CTNNB1* whereas Huh7 have a *TP53* mutation. In patients, β-catenin and p53 mutations are usually mutually exclusive and may represent different subtypes of liver cancer.

From our CLIP-seq we observed that TRIM71 was binding to the RNAs of *TCF7* and *TCF7L2*, two transcription factors that bind to activated β-catenin and are required for β-catenin associated transcription. We further showed that TRIM71 was positively regulating the RNAs and depletion of *TRIM71* decreased their protein levels. We used a TOPflash reporter which is activated by β-catenin/TCF binding and found that knockdown of *TRIM71* had a small and inconsistent effect on reporter activity. We next tested whether overexpression of *TRIM71* could potentiate activation of TOPflash and we again saw a small effect. Thus, it is possible that the importance of TRIM71 in liver cancer cells could be specific to the subtypes where there is a mutation in the Wnt pathway leading to constitutive β-catenin associated transcription. However, given the small and inconsistent effect associated with *TRIM71* perturbation, it is not completely clear.

Future perspectives

Through this work we have shown that the roles of TRIM71 in RNA regulation are broader than previously thought. TRIM71 does not only function to inhibit RNAs. In fact, there may be more RNAs promoted by TRIM71 than inhibited. Further, the interaction of TRIM71 with the CRD complex members IMP1 and DHX9 provide insights into potential ways that TRIM71 might positively regulating RNAs. Although we did not touch on

potential mechanisms in this thesis, it would be important to look at TRIM71 in the context of its protein binding partners. IMP1 is a fairly well studied RNA binding protein that plays multiple regulatory roles. It is capable of stabilizing RNAs such as *MYC* by preventing degradation by nucleases. IMP1 has be shown to recognize m6A RNA modifications and shuttle RNAs into stress granules during high stress conditions for the cell. Although we have not studied TRIM71 in the context of stress conditions it would be interesting to see whether TRIM71 shares any mechanistic similarities with IMP1.

Despite the binding of TRIM71 to many oncogenes in the CLIP, the role of TRIM71 in cancer seems small and it may in fact be a bystander gene resulting from expression of *LIN28B*. However, it was apparent that cancer cell lines do not represent the full heterogeneity in patients. AML cell lines did not mimic the high *TRIM71* expression seen in patients in The Cancer Genome Atlas database. For this reason, following up on AML did not prove fruitful. Future studies on TRIM71 in cancer may focus on *TRIM71* overexpression mouse models, or patient samples, where the expression of *TRIM71* may be higher than it is in currently available cancer cell lines.

Finally, in the process of carrying out this work, exciting new research revealed that *TRIM71* has a role in the developmental brain disorder congenital hydrocephalus. This disease is intriguing as it relates to neural defects seen in *Trim71* mutant mice. Furthermore, the disease appears to be dependent on the RNA binding abilities of TRIM71. We have shown here that TRIM71 is capable of binding thousands of potential RNAs. Some of these newly elucidated targets such as the *MBNL* and *NFI* families have been strongly associated with neural development. Following up on these targets in a more relevant cellular context such as neural progenitor cells could be important and consequential for human health.

Chapter 5 - Materials and Methods

Cell Culture

Hep3B and HepG2 cells were purchased from the American Type Culture Collections (ATCC). Huh-7 and Huh-6 cells were purchased from the Japanese Collection of Research Bioresources (JCRB) Cell Bank. All cell lines were cultured in media containing 10% Fetal Bovine Serum (FBS), penicillin, and streptomycin. HepG2 were cultured in Dulbecco's Modified Eagles Medium (DMEM) with 4.5g/L glucose. Huh-7 cells were grown in DMEM with 1g/L glucose. Hep3B were grown in Eagle's Minimal Essential Media. Huh6 cells were grown in RPMI with 4% FBS.

RNA extraction, Reverse Transcription, and Quantitative PCR

Cells were lysed in TRIzol reagent and RNA was isolated using the Direct-zol RNA Kit (Zymo). For CLIP experiments we used the microprep kit whereas all other RNA was isolated using the miniprep plus kit.

First strand cDNA synthesis was performed using Superscript IV Reserve Transcriptase (Thermo Fischer cat. # 18091050). For CLIP experiments, cDNA was primed using random hexamers. For all other experiments, oligo dT was used. For expression analysis, 1ug of total RNA was reverse transcribed. For CLIP and RIP

experiments, all eluted RNA was loaded into the RT reaction for IP samples and 1-3uL of RNA from inputs were used.

All qPCR was SYBR Green based and used either LightCycler FastStart DNA Master SYBR Green I (Roche) or QuantiTech SYBR Green (Qiagen). Roche was used for expression analysis. Qiagen was used for CLIP and RIP experiments. A master mix was made for each primer reaction and run in technical triplicates using a LightCycler 480 System (Roche).

Antibodies

Used for western blot: alpha-Tubulin (#2144), c-MYC (#5605), LIN28B (#4196), p21 (#2947), EGR1 (#4154), IMP1 (#8482), HA-tag (#3724) from Cell Signaling Technology; LIN-41 (55003-1-AP), p16INK4A (10883-1-AP), MBNL3 (24610-1-AP), NFIA (11750-1-AP) from Proteintech. MBNL1 (sc-47740) from Santa Cruz Biotechnology.

Used for IP and cross-linking IP: IMP1 (#8482) from Cell Signaling Technology; HA-tag (H6908) and IgG (R5506) from Sigma-Aldrich.

Western Blotting

Cells on 10cm plates were lysed in 1mL of RIPA buffer, added to a microcentrifuge tube, and incubated on ice for 10 minutes and vortexed every 2 minutes. Cell lysates was centrifuged at 13,200rpm for 20 minutes and supernatant was added to a fresh microcentrifuge tube. Laemmli buffer was added to the cleared lysates and proteins were size separated using 4-12% NuPAGE gels (Thermo Fisher #NP0316BOX) with corresponding buffer (Thermo Fisher #NP0001). Proteins were transferred to nitrocellulose or PVDF membranes and blocked in 3-5% milk. Primary antibodies were

incubated at 4°C overnight. Membranes were washed 6 times with TBST, with the final 3 washes rocking for 5 minutes each. HRP-conjugated secondary antibodies (Santa Cruz) were incubated at room temperature for 2 hours at a 1:10,000 dilution. Membranes were washed 6 times with TBST, with the final 3 washes rocking for 5 minutes each. Protein was detected using Chemiluminescent substrate (Thermo Fisher #34580).

Quantification of westerns

Western bands were quantified using ImageStudioLite (Licor version 5.2.5). Test genes were normalized to the loading control by dividing the two signals (Test ÷ Control). This normalized signal was calculated for each treatment conditions (e.g. siNEG, etc.). Next, siRNA treatment conditions were compared by dividing the gene specific siRNA samples by siNEG (e.g. normalized siTRIM71 ÷ normalized siNEG). Thus, values greater than 0 indicate genes increased with treatment and values less than 0 indicates genes decreased with treatment.

siRNA transfections

Cells were transfected with Thermo Fischer Scientific pre-designed Silencer Select siRNA, cat. # 4392420 and 4390843 at a final concentration of 20nM. Assay IDs for the siRNAs used for each gene are as follows: TRIM71 (s43598, s43599), IGF2BP1 (s20916, s20917), cMYC (s9129, s9130), EGR1 (s4537, s4539), LIN28B (s52477, s52478, s52479), CDKN1A (s416), CTNNB1 (s438).

For siRNA knockdown, Silencer Select Pre-designed siRNA and Negative Control #1 siRNA were purchased from Thermo Fischer Scientific. The target sequences for TRIM71 are as follows: TRIM71 siRNA #1 (s43598), 5'-

GAAUGAUCGUUGUGGUGGATT-3'; TRIM71 siRNA #2 (s43599), 5'-GGACAUACGUGGUGAGUUATT-3'.

All siRNA transfections were performed using Lipfectamine RNAiMax cat. # 13778150. Transfections were optimized for each cell line to obtain the best mRNA knockdown efficiency. HepG2 cells were reverse transfected in 6-well plates for 3 days before being split for further cell assays. Huh6 cells were reverse transfected for 3 days in 96-well plates and assays were performed without splitting the cells. Huh7 and Hep3B cells were forward transfected for 3 days in 6-well plates before being split for further cell assays.

MTT Assay

Cells were depleted for the indicated genes using siRNAs and were seeded for MTT assay on day 4, 1 day following the standard 3-day transfection. HepG2 cells were seeded 1,000 cells/well in a 96-well plate in technical quadruplicate with 100uL of media per well. Huh7 cells were seeded 2,000 cells/well. For each day of the assay, MTT dye was added to the cells, incubated for 2 hours, aspirated, dissolved in 100uL DMSO, and then the absorbance was read at 565nm.

BrdU Assay

Millipore BrdU Cell Proliferation Assay kit cat. # 2750 was used to detect and measure BrdU incorporation. Briefly, 20,000 cells/well were seeded into 96-well plates in triplicate after the standard 3-day siRNA transfection. BrdU reagent was added to the cells as instructed and incubated for 18 hours, before proceeding with the kit protocol.

Senescence staining

The Cell Signaling Technologies Senescence Associated β-galactosidase staining kit cat. # 9860 was used following the suggested protocol. The staining reagent was adjusted to a pH of 6 for every experiment. Cell were typically cultured for 7 days following the standard 3-day transfection protocol to allow time for the cells to become senescent. The siRNA negative treated cells needed to be split during this time to prevent confluence. Otherwise confluence would lead to falsely positive senescence staining. Cells were counted under a light microscope as either positive or negative for staining.

Cross-linking IP and qRT-PCR

Cross-linking IP followed by qRT-PCR: We used a CLIP protocol modified from Moore, et al (Moore et al., 2016). Cells were incubated for 14-18 hours in media supplemented with 4-SU at a final concentration of 100uM. Cells were washed twice with 10mL of cold PBS and 4mL of cold PBS was added to the cells.

Cells overlaid with PBS were crosslinked for 72 seconds with a Stratalinker 1800 emitting 3000uL/sec/cm2 at 365nM. Cells were scraped into a 15mL conical tube then centrifuged at 1200rpm for 6 min. PBS was aspirated and cells were resuspended in ~1mL of cold PBS. Cells were transferred to a microcentrifuge tube and pelleted at 1000G for 5 min. Supernatant was aspirated and cell pellet was frozen at -80°C for later use.

Buffers were made following the Moore, *et al*., 2014 protocol at least 1 day but not more than 1 month before the IP protocol. Beads were prepared for IP one day before the CLIP protocol was carried out. Protein G dynabeads (Thermo Fisher #10004D) were

washed 3x with bead wash buffer (BWB) and resuspended in 500uL of BWB. 3ug of anti-HA antibody (Sigma #H6908) was added and beads were rotated overnight at 4°C. The next day the beads were washed 3x with PXL buffer and kept on ice in a minimal volume of PXL until ready to use.

On day 2, frozen cell pellet was resuspended in 1mL of 1x PXL buffer supplemented with HALT protease inhibitor cocktail (Thermo Fisher #78420). 30uL of Superase-In (600 units of enzyme) (Invitrogen #AM2696) and 15uL DNaseI (30 units of enzyme) (New England Biolabs #M0303L) was added. Cells were placed on ice for 10 minutes and vortexed every 2 minutes. Cells were then thermomixed for 20min at 37°C at 1000rpm. Samples were centrifuged at 13,200rpm in 4°C for 25 minutes. 1mL of supernatant was added to a new microcentrifuge tube. 50uL (5%) of supernatant was taken as an input control. The remaining 950uL of supernatant was added to protein G dynabeads prepared as described above. Tubes were rotated end over end for 2 hours at 4°C.

During the 2-hour incubated, diluted proteinase K was prepared as follows. The 20mg/mL stock of proteinase K (New England Biolabs #P8107S) was diluted to 4mg/uL in PK buffer. 200uL was prepared for each IP sample and 100uL was prepared for each input sample. After diluting proteinase K, it was incubated in a thermomixer for 20 minutes at 37°C at 1000rpm.

After incubating the beads with the cell lysate, CLIP washes were performed as follows from the buffers prepared according to Moore, *et al*., 2014. The stringent washes were as follows with 1mL of buffer used per wash: 3x washes of 1x PXL cell lysis buffer, 3x washes of high salt buffer, 3x washes of high stringency buffer, 3x washes of low salt

buffer, 2x washes of PNK buffer. After aspiration of the final wash, 200uL and 100uL of diluted proteinase K solution was added to each IP and input sample, respectively. The beads with proteinase K were incubated in a thermomixer for 20 minutes at 37°C at 1000rpm. Next, 600uL of TRIzol was added to each sample. Note that TRIzol requires a 3:1 ratio when added to a liquid sample so 600uL was sufficient for a 200uL liquid sample. From here, RNA was extracted using a ZYMO DirectZol microkit as described above.

Samples were eluted from the Zymo DirectZol column with 15uL of water. For qPCR experiments, cDNA was synthesized as described above. For CLIP-sequencing, RNA samples were submitted to the Dana Farber Molecular Biology Core Facility.

Crosslinking co-IP

For crosslinking co-IP experiments carried out in the presence of RNase, 5uL of RNase was added at the same time as DNase, instead of SuperaseIn, before the CLIP washes. Centrifugation, immunoprecipitation, and the following washes were then carried out as described above. After the final CLIP wash, beads were resuspended in 60uL of PNK buffer. For all crosslinking co-IP samples, an additional 1uL of RNase was added to the IP samples after the CLIP washes and 2uL of RNase was added to input samples. These were thermomixed for 20 minutes at 37°C at 1000rpm. From here, 30uL of 4x Laemmli buffer was added to the beads and 50uL of 2x Laemmli buffer was added to input samples. Samples were then incubated at 95°C for 5 minutes.

CLIP-sequencing

Library preparation and sequencing was performed by the Dana Farber Molecular Biology Core facility.

RNA quality was analyzed using a bioanalyzer. RNA integrity numbers (RIN) were low for input samples, between 2-3, due to incubation of cell lysate with proteinase K. Despite the slightly degraded RNA input samples, there were no indications of bad quality in mapping of the RNA. RIN values for the IP samples ranged from 2-7, however the RIN values are not necessarily indicative of good RNA quality of IP samples.

cDNA libraries were prepared using the SMARTer V4 kit for low input total RNA (Takara #635007) which uses random hexamer priming and depletion of ribosomal RNA. Despite high concentration of RNA in the input samples, the same kit was used for input as well as IP samples to avoid bias in library prep. A total of 2ng of RNA was used for library prep of each sample.

Libraries were sequenced on an Illumina NextSeq500 with single end reads of 75bp in length. All 20 samples were sequenced in the same sequencing run to an average depth of ~21 million reads with the range between 18-30 millions reads. IP samples generally had fewer reads than the input samples. Ribosomal RNA represented around 5% of the IP reads and 10% of input reads indicating successful ribosomal RNA depletion. Reads were aligned to the genome using bowtie and PCR duplicates were flagged and removed using bamsormadup and samtools. Differential expression analysis was performed using DEseq2. The PAR-CLIP data analyzer (Corcoran et al., 2011) was used to calculate T to C transition fraction for each gene.

RNA IP

For RIP experiments, cells were grown in 10cm plates to near confluence. Cells were washed with cold PBS and lysed in 1mL of cell lysis buffer (Cell Signaling Technologies cat. #9803) supplemented with HALT protease inhibitor (Thermo Fisher

#78430) and Superase-In (Invitrogen #AM2696). Lysates were cleared by centrifugation at 14,000g for 10min. 50uL of lysate was taken as an input control and 300uL of TRIzol was added then frozen at -20°C. 5ug of antibody was added to the remaining lysate and incubated rotating overnight at 4°C.

The next day, Protein G dynabeads were added and incubated for 2 hours at 4°C. Beads were washed 5x with cell lysis buffer. After washes, beads were eluted in 350uL of TRIzol. RNA extraction proceeded as above.

Plasmids and cloning

HA-TRIM71 pMXs plasmids and deletion mutants were obtained from Addgene (item #s: 52717, 52718, 52719, 52720, 52721, 52722, 52726). *LIN28B*, *MYC*, *GFP*, and *MDM2* pcDNA3 plasmids for overexpression experiments were ordered from Addgene (item #s: 51373, 74162, 13031, 16233, respectively).

TRIM71 and other overexpression gene sequences were subcloned from pMXs and pcDNA3 plasmids into a puromycin selectable pLVX plasmid (Clontech #632159) for lentiviral based stable transductions. For transductions we used Addgene plasmids psPAX2 (#122560) and pCMV-VSV-G (#8454).

TOPflash plasmids and a control renilla luciferase plasmid were ordered from Addgene (item #s: 12456, 12457, 27163).

Page intentionally left blank

Bibliography

Abrahante, J.E., Miller, E.A., and Rougvie, A.E. (1998). Identification of heterochronic mutants in Caenorhabditis elegans: temporal misexpression of a collagen:: green fluorescent protein fusion gene. Genetics *149*, 1335-1351.

Aguiar, T.F.M., and Carneiro, T.N. (2017). The genetic and epigenetic landscapes of hepatoblastomas. Applied Cancer ….

Balzeau, J., Menezes, M.R., Cao, S., and in genetics, H.-J.P. (2017). The LIN28/let-7 pathway in cancer. Frontiers in genetics.

Barry, G., Piper, M., Lindwall, C., and of …, M.-R. (2008). Specific glial populations regulate hippocampal morphogenesis. Journal of ….

Bohlman, S., and Manfredi, J.J. (2014). Mutant p53 and MDM2 in Cancer. Sub-cellular biochemistry *85*, 235-246.

Buchan, J.R. (2014). mRNP granules: assembly, function, and connections with disease. RNA biology.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell *150*, 1209-1222.

Büssing, I., Slack, F.J., and Großhans, H. (2008). let-7 microRNAs in development, stem cells and cancer. Trends in Molecular Medicine *14*, 400-409.

Chang, H.-M., Martinez, N.J., Thornton, J.E., Hagan, J.P., Nguyen, K.D., and Gregory, R.I. (2012). Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nature Communications *3*, 923.

Charizanis, K., Lee, K.Y., Batra, R., Goodwin, M., and Neuron, Z.-C. (2012). Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron.

Chen, J., Lai, F., and Niswander, L. (2012). The ubiquitin ligase mLin41 temporally promotes neural progenitor cell maintenance through FGF signaling. Genes & Development *26*, 803-815.

Chen, Y.L., Yuan, R.H., Yang, W.C., Hsu, H.C., and Jeng, Y.M. (2013). The stem cell E3-ligase Lin-41 promotes liver cancer progression through inhibition of microRNA-mediated gene silencing. The Journal of Pathology *229*, 486-496.

Cho, J., Chang, H., Kwon, C.S., Kim, B., Kim, Y., Choe, J., Ha, M., Kim, Y., and Kim, N.V. (2012). LIN28A is a suppressor of ER-associated translation in embryonic stem cells. Cell *151*, 765-777.

Conway, A.E., Van Nostrand, E.L., Pratt, G.A., Aigner, S., Wilbert, M.L., Sundararaman, B., Freese, P., Lambert, N.J., Sathe, S., Liang, T.Y.*, et al.* (2016). Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. Cell Reports *15*, 666-679.

Coppé, J.-P., Desprez, P.-Y., Krtolica, A., and Campisi, J. (2010). The senescence-associated secretory phenotype: the dark side of tumor suppression. Annual Review of Pathological Mechanical Disease *5*, 99-118.

Corcoran, D.L., Georgiev, S., Mukherjee, N., Gottwein, E., Skalsky, R.L., Keene, J.D., and Ohler, U. (2011). PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. Genome Biol *12*, R79.

Degrauwe, N., Schlumpf, T.B., Janiszewska, M., Martin, P., Cauderay, A., Provero, P., Riggi, N., Suvà, M.-L.L., Paro, R., and Stamenkovic, I. (2016a). The RNA Binding Protein IMP2 Preserves Glioblastoma Stem Cells by Preventing let-7 Target Gene Silencing. Cell reports *15*, 1634-1647.

Degrauwe, N., Suvà, M.-L., Janiszewska, M., Riggi, N., and Stamenkovic, I. (2016b). IMPs: an RNA-binding protein family that provides a link between stem cell maintenance in normal development and cancer. Genes & Development *30*, 2459-2474.

Doyle, G.A.R., Leeds, P.F., Fleisig, A.J., and acids …, R.J. (1998). The c-myc coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins. The c-myc coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins.

Edelmann, S., Fahrner, R., Malinka, T., and Liver …, S.-B.H. (2015). Nuclear Factor I-C acts as a regulator of hepatocyte proliferation at the onset of liver regeneration. Liver ….

Elcheva, I., Goswami, S., Noubissi, F.K., and Spiegelman, V.S. (2009). CRD-BP Protects the Coding Region of βTrCP1 mRNA from miR-183-Mediated Degradation. Molecular Cell *35*, 240-246.

Furey, C., Choi, J., Jin, S., Zeng, X., Timberlake, A.T., Nelson-Williams, C., Mansuri, S.M., Lu, Q., Duran, D., Panchagnula, S.*, et al.* (2018). De Novo Mutation in Genes

Regulating Neural Stem Cell Fate in Human Congenital Hydrocephalus. Neuron *99*, 302-3140000.

Gao, S., and Wang, Z. (2012). Subcellular localization of p44/WDR77 determines proliferation and differentiation of prostate epithelial cells. PloS one.

Goldman, M., Craft, B., Hastie, M., Repečka, K., Kamath, A., McDade, F., Rogers, D., Brooks, A.N., Zhu, J., and Haussler, D. (2019). The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. bioRxiv, 326470.

Graf, R., Munschauer, M., Mastrobuoni, G., and Mayr, F. (2013). Identification of LIN28B-bound mRNAs reveals features of target recognition and regulation. Identification of LIN28B-bound mRNAs reveals features of target recognition and regulation.

Gronostajski, R.M. (2000). Roles of the NFI/CTF gene family in transcription and development. Gene.

Großhans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. Developmental Cell *8*, 321-330.

Guichard, C., Amaddeo, G., Imbeaud, S., and Nature …, L.-Y. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nature ….

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C.C., Munschauer, M.*, et al.* (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell *141*, 129-141.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nature Structural & Molecular Biology *16*, 1021-1025.

Hansen, T.V.O., Hammer, N.A., Nielsen, J., Madsen, M., Dalbaeck, C., Wewer, U.M., Christiansen, J., and Nielsen, F.C. (2004). Dwarfism and Impaired Gut Development in Insulin-Like Growth Factor II mRNA-Binding Protein 1-Deficient Mice. Molecular and Cellular Biology *24*, 4448.

Hatakeyama, S. (2011). TRIM proteins and cancer. Nature Reviews Cancer *11*, 792-804.

Hatakeyama, S. (2017). TRIM family proteins: roles in autophagy, immunity, and carcinogenesis. Trends in biochemical sciences *42*, 297-311.

Heng, Y.H.E., Barry, G., Richards, L.J., and Neurosignals, P.-M. (2012). Nuclear factor I genes regulate neuronal migration. Neurosignals.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, N.V. (2008). Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. Molecular Cell *32*, 276-284.

Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B., Mesquita, A., Liu, C., Yuan, C.L.*, et al.* (2018). Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nature Cell Biology *20*, 285-295.

Jiang, S., and Baltimore, D. (2016). RNA-binding protein Lin28 in cancer and immunity. Cancer Letters *375*, 108-113.

Johnson, C.D., Esquela-Kerscher, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., Wilson, M., Wang, X., Shelton, J., Shingara, J.*, et al.* (2007). The let-7 microRNA represses cell proliferation pathways in human cells. Cancer research *67*, 7713-7722.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS Is Regulated by the let-7 MicroRNA Family. Cell *120*, 635-647.

Johnson, S.M., Lin, S.-Y., and Slack, F.J. (2003). The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. Developmental Biology *259*, 364-379.

Kim, T.H., Bae, C.H., Yang, S., and cell …, P.-J.C. (2015). Nfic regulates tooth root patterning and growth. Anatomy & cell ….

Klijn, C., Durinck, S., Stawiski, E.W., Haverty, P.M., Jiang, Z., Liu, H., Degenhardt, J., Mayba, O., Gnad, F., Liu, J.*, et al.* (2015). A comprehensive transcriptional portrait of human cancer cell lines. Nature Biotechnology *33*, 306-312.

Koch, A., Denkhaus, D., Albrecht, S., and research, L.-I. (1999). Childhood hepatoblastomas frequently carry a mutated degradation targeting box of the β-catenin gene. Cancer research.

Konieczny, P., and Stepniak-Konieczna, E. (2014). MBNL proteins and their target RNAs, interaction and splicing regulation. Nucleic acids ….

Kumari, P., Aeschimann, F., Gaidatzis, D., Keusch, J.J., Ghosh, P., Neagu, A., Pachulska-Wieczorek, K., Bujnicki, J.M., Gut, H., Großhans, H.*, et al.* (2018). Evolutionary plasticity of the NHL domain underlies distinct solutions to RNA recognition. Nature Communications *9*, 1549.

Lancman, J.J., Caruccio, N.C., Harfe, B.D., Pasquinelli, A.E., Schageman, J.J., Pertsemlidis, A., and Fallon, J.F. (2005). Analysis of the regulation of lin-41 during chick and mouse limb development. Developmental dynamics : an official publication of the American Association of Anatomists *234*, 948-960.

Lazakovitch, E., and Kalb, J.M. (2005). nfi-1 affects behavior and life-span in C. elegans but is not essential for DNA replication or survival. BMC ….

Lee, J.E., and Cooper, T.A. (2009). Pathogenic mechanisms of myotonic dystrophy. Pathogenic mechanisms of myotonic dystrophy.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, S., Cho, S., Kim, S.M., Choi, K., Cho, J., Gwak, H.-S., Kim, Y.-J., Yoo, H., Lee, S.-H., Park, J.*, et al.* (2014). The ubiquitin ligase human TRIM71 regulates let-7 microRNA biogenesis via modulation of Lin28B protein. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1839*, 374-386.

Lehrbach, N.J., Armisen, J., Lightfoot, H.L., Murfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nature Structural & Molecular Biology *16*, 1016-1020.

Lemm, I., and Ross, J. (2002). Regulation of c-myc mRNA decay by translational pausing in a coding region instability determinant. Molecular and cellular biology.

Li, Y.-P., Duan, F.-F., Zhao, Y.-T., Gu, K.-L., Liao, L.-Q., Su, H.-B., Hao, J., Zhang, K., Yang, N., and Wang, Y. (2019). A TRIM71 binding long noncoding RNA Trincr1 represses FGF/ERK signaling in embryonic stem cells. Nature Communications *10*, 1368.

Lin, Y.-C., Hsieh, L.-C., Kuo, M.-W., Yu, J., Kuo, H.-H., Lo, W.-L., Lin, R.-J., Yu, A.L., and Li, W.-H. (2007). Human TRIM71 and Its Nematode Homologue Are Targets of let-7 MicroRNA and Its Zebrafish Orthologue Is Essential for Development. Molecular Biology and Evolution *24*, 2525-2534.

Loedige, I., Gaidatzis, D., Sack, R., Meister, G., and Filipowicz, W. (2013). The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic Acids Research *41*, 518-532.

Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R.*, et al.* (2015). The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. Cell Reports *13*, 1206-1220.

Lu, W., Quintero-Rivera, F., Fan, Y., and PloS …, A.-F.S. (2007). NFIA haploinsufficiency is associated with a CNS malformation syndrome and urinary tract defects. PLoS ….

Mackereth, C.D., and Sattler, M. (2012). Dynamics in multi-domain protein recognition of RNA. Current opinion in structural biology.

Mansfield, J.H., Harfe, B.D., Nissen, R., Obenauer, J., Srineel, J., Chaudhuri, A., Farzan-Kashani, R., Zuker, M., Pasquinelli, A.E., and Ruvkun, G. (2004). MicroRNA-responsive'sensor'transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nature genetics *36*, 1079.

Martynoga, B., Mateo, J.L., and …, Z.-B. (2013). Epigenomic enhancer annotation reveals a key role for NFIX in neural stem cell quiescence. Genes & ….

Mayr, C., Hemann, M.T., and Science, B.-D.P. (2007). Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. Science.

Meyers, R.M., Bryan, J.G., McFarland, J.M., Weir, B.A., Sizemore, A.E., Xu, H., Dharia, N.V., Montgomery, P.G., Cowley, G.S., Pantel, S.*, et al.* (2017). Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. Nature Genetics *49*, 1779.

Mihara, K., Miyazaki, M., and Culture …, F.-K. (1997). The p53 gene status and other cellular characteristics of human cell lines maintained in our laboratory. TISSUE CULTURE ….

Mitschka, S., Ulas, T., Goller, T., Schneider, K., Egert, A., Mertens, J., Brüstle, O., Schorle, H., Beyer, M., Klee, K.*, et al.* (2015). Co-existence of intact stemness and priming of neural differentiation programs in mES cells lacking Trim71. Scientific Reports *5*.

Molenaar, J.J., Domingo-Fernández, R., Ebus, M.E., Lindner, S., Koster, J., Drabek, K., Mestdagh, P., van Sluis, P., Valentijn, L.J., van Nes, J.*, et al.* (2012). LIN28B induces neuroblastoma and enhances MYCN levels via let-7 suppression. Nature Genetics *44*, 1199.

Moore, M.J., Zhang, C., Gantman, E., Mele, A., Darnell, J.C., and Darnell, R.B. (2016). Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. Nature Protocols *9*, 263-293.

Müller, S., Bley, N., Glaß, M., Busch, B., Rousseau, V., Misiak, D., Fuchs, T., Lederer, M., and Hüttelmaier, S. (2018). IGF2BP1 enhances an aggressive tumor cell phenotype by impairing miRNA-directed downregulation of oncogenic factors. Nucleic Acids Research.

Müller-McNicoll, M., and Neugebauer, K.M. (2013). How cells get the message: dynamic assembly and function of mRNA–protein complexes. Nature Reviews Genetics *14*, 275.

Murray, M.J., Saini, H.K., Siegler, C.A., Hanning, J.E., Barker, E.M., van Dongen, S., Ward, D.M., Raby, K.L., Groves, I.J., Scarpini, C.G.*, et al.* (2013). LIN28 Expression in Malignant Germ Cell Tumors Downregulates let-7 and Increases Oncogene Levels. Cancer Research *73*, 4872-4884.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Cell, S.-P. (2011). Molecular basis for interaction of let-7 microRNAs with Lin28. Cell.

Nguyen, D., Richter, D., Michel, G., Mitschka, S., Kolanus, W., Cuevas, E., and Wulczyn, G.F. (2017). The ubiquitin ligase LIN41/TRIM71 targets p53 to antagonize cell death and differentiation pathways during stem cell differentiation. Cell Death & Differentiation *24*, 1063-1078.

Nguyen, L.H., Robinton, D.A., Seligson, M.T., Wu, L., Li, L., Rakheja, D., Comerford, S.A., Ramezani, S., Sun, X., Parikh, M.S*., et al.* (2014). Lin28b Is Sufficient to Drive Liver Cancer and Necessary for Its Maintenance in Murine Models. Cancer Cell *26*, 248-261.

Nishino, J., Kim, S., Zhu, Y., Zhu, H., and Morrison, S.J. (2013). A network of heterochronic genes including Imp1 regulates temporal changes in stem cell properties. eLife *2*.

Noubissi, F.K., Elcheva, I., Bhatia, N., Shakoori, A., Ougolkov, A., Liu, J., Minamoto, T., Ross, J., Fuchs, S.Y., and Spiegelman, V.S. (2006). CRD-BP mediates stabilization of βTrCP1 and c-myc mRNA in response to β-catenin signalling. Nature *441*.

Noubissi, F.K., Goswami, S., Sanek, N.A., and research, K.K. (2009). Wnt signaling stimulates transcriptional outcome of the Hedgehog pathway by stabilizing GLI1 mRNA. Wnt signaling stimulates transcriptional outcome of the Hedgehog pathway by stabilizing GLI1 mRNA.

Ohtani, N., Yamakoshi, K., and of Medical …, T.-A. (2004). The p16INK4a-RB pathway: molecular link between cellular senescence and tumor suppression. The Journal of Medical ….

Pasquinelli, A.E., Reinhart, B.J., Slack, F., and Nature, M.-M.Q. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature.

Piskounova, E., Polytarchou, C., Thornton, J.E., and Gregory, R.I. (2011). Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. Cell.

Plachez, C., Lindwall, C., Sunn, N., and of …, P.-M. (2008). Nuclear factor I gene expression in the developing forebrain. Journal of ….

Qian, Y., and Senescence, C.-X. (2013). Senescence regulation by the p53 protein family. Cell Senescence.

Rand, T.A., Sutou, K., Tanabe, K., Jeong, D., Nomura, M., Kitaoka, F., Tomoda, E., Narita, M., Nakamura, M., and Nakamura, M. (2018). MYC Releases Early Reprogrammed Human Cells from Proliferation Pause via Retinoblastoma Protein Inhibition. Cell reports *23*, 361-375.

Ray, D., Ha, K., Nie, K., Zheng, H., Hughes, T.R., and Morris, Q.D. (2017). RNAcompete methodology and application to determine sequence preferences of unconventional RNA-binding proteins. Methods *118*, 3-15.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, R.H., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.

Ren, H., Xu, Y., Wang, Q., Jiang, J., Wudumuli, Hui, L., Zhang, Q., Zhang, X., Wang, E., Sun, L.*, et al.* (2017). E3 ubiquitin ligase tripartite motif-containing 71 promotes the proliferation of non-small cell lung cancer through the inhibitor of kappaB-α/nuclear factor kappaB pathway. Oncotarget *9*, 10880-10890.

Roush, S., and Slack, F.J. (2008). The let-7 family of microRNAs. Trends in cell biology *18*, 505-516.

Rufini, A., Tucci, P., Celardo, I., and Oncogene, M.-G. (2013). Senescence and aging: the critical roles of p53. Oncogene.

Rupaimoole, R., and Slack, F.J. (2017). MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nature Reviews Drug Discovery *16*, 203-222.

Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E.A., Michel, G., Nitsch, R., Krappmann, D., and Wulczyn, G.F. (2009). The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. Nature Cell Biology *11*.

Schulman, B.R., Esquela-Kerscher, A., and Slack, F.J. (2005). Reciprocal expression of lin-41 and the microRNAs let-7 and mir-125 during mouse embryogenesis. Developmental Dynamics *234*, 1046-1054.

Schulman, B.R., Liang, X., Stahlhut, C., DelConte, C., Stefani, G., and Slack, F.J. (2008). The let-7 microRNA target gene, Mlin41/Trim71 is required for mouse embryonic survival and neural tube closure. Cell cycle (Georgetown, Tex) *7*, 3935-3942.

Schulze, K., Imbeaud, S., Letouzé, E., and Nature …, A.-L.B. (2015). Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nature ….

Shyh-Chang, N., and Daley, G.Q. (2013). Lin28: primal regulator of growth and metabolism in stem cells. Cell stem cell *12*, 395-406.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (1999). The lin-41 RBCC Gene Acts in the C. elegans Heterochronic Pathway between the let-7 Regulatory RNA and the LIN-29 Transcription Factor. Molecular Cell *5*.

Slack, F.J., and Ruvkun, G. (1998). A novel repeat domain that is often associated with RING finger and B-box motifs. Trends in Biochemical Sciences *23*, 474475.

Sparanese, D., and Lee, C.H. (2007). CRD-BP shields c-myc and MDR-1 RNA from endonucleolytic attack by a mammalian endoribonuclease. Nucleic Acids Research *35*, 1209-1221.

Spilker, K.A., and Wang, G.J. (2012). Caenorhabditis elegans Muscleblind homolog mbl-1 functions in neurons to regulate synapse formation. Neural ….

Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. Nature Reviews Molecular Cell Biology *9*, 219-230.

Stöhr, N., Köhn, M., Lederer, M., and …, G.-M. (2012). IGF2BP1 promotes cell migration by regulating MK5 and PTEN signaling. Genes & ….

Stopa, N., Krebs, J.E., and and molecular life sciences, S.-D. (2015). The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. Cellular and molecular life sciences.

Sznajder, Ł., Michalak, M., and acids …, T.-K. (2016). Mechanistic determinants of MBNL activity. Nucleic acids ….

Thornton, J.E., Chang, H.-M., Piskounova, E., and Gregory, R.I. (2012). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). Rna *18*, 1875-1885.

Thornton, J.E., and Gregory, R.I. (2012). How does Lin28 let-7 control development and disease? Trends in Cell Biology *22*, 474-482.

Tocchini, C., and Ciosk, R. (2015). TRIM-NHL proteins in development and disease. Seminars in Cell & Developmental Biology *47*, 52-59.

Tocchini, C., Keusch, J.J., Miller, S.B., Finger, S., Gut, H., Stadler, M.B., and Ciosk, R. (2014). The TRIM-NHL Protein LIN-41 Controls the Onset of Developmental Plasticity in Caenorhabditis elegans. PLoS Genetics *10*.

Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M.*, et al.* (2017). Defining a Cancer Dependency Map. Cell *170*, 564-1916796928.

Tu, H.-C., Schwitalla, S., Qian, Z., LaPier, G.S., Yermalovich, A., Ku, Y.-C., Chen, S.-C., Viswanathan, S.R., Zhu, H., Nishihara, R.*, et al.* (2015). LIN28 cooperates with WNT signaling to drive invasive intestinal and colorectal adenocarcinoma in mice and humans. Genes & Development *29*, 1074-1086.

Urbach, A., Yermalovich, A., Zhang, J., Spina, C.S., Zhu, H., Perez-Atayde, A.R., Shukrun, R., Charlton, J., Sebire, N., Mifsud, W.*, et al.* (2014). Lin28 sustains early renal progenitors and induces Wilms tumor. Genes & Development *28*, 971-982.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K.*, et al.* (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nature Methods *13*.

Vella, M.C., Choi, E.-Y., Lin, S.-Y., Reinert, K., and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. Genes & Development.

Viswanathan, S.R., and Daley, G.Q. (2010). Lin28: A MicroRNA Regulator with a Macro Role. Cell *140*, 445-449.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97-100.

Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L.*, et al.* (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nature Genetics *41*, 843-848.

Wade, M., Li, Y.C., and Cancer, W.-G.M. (2013). MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nature Reviews Cancer.

Wang, J., Zheng, T., Chen, X., and of …, S.-X. (2011). MDM2 antagonist can inhibit tumor growth in hepatocellular carcinoma with different types of p53 in vitro. Journal of ….

Wei, W., Chua, M.S., and journal of …, G.-S. (2010). Small molecule antagonists of Tcf4/β-catenin complex inhibit the growth of HCC cells in vitro and in vivo. International journal of ….

Weidensdorfer, D., Stöhr, N., Baude, A., Lederer, M., Köhn, M., Schierhorn, A., Buchmeier, S., Wahle, E., and Hüttelmaier, S. (2009). Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. RNA *15*, 104-115.

Wilbert, M.L., Huelga, S.C., Kapeli, K., Stark, T.J., Liang, T.Y., Chen, S.X., Yan, B.Y., Nathanson, J.L., Hutt, K.R., Lovci, M.T.*, et al.* (2012). LIN28 Binds Messenger RNAs at GGAGA Motifs and Regulates Splicing Factor Abundance. Molecular cell *48*, 195-206.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nature Cell Biology *11*, 228-234.

Worringer, K.A., Rand, T.A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., Narita, M., Srivastava, D., and Yamanaka, S. (2014). The let-7/LIN-41 Pathway Regulates Reprogramming to Human Induced Pluripotent Stem Cells by Controlling Expression of Prodifferentiation Genes. Cell Stem Cell *14*, 40-52.

Yin, J., Kim, T.-H., Park, N., Shin, D., Choi, H., Cho, S., Park, J., and Kim, J. (2016). TRIM71 suppresses tumorigenesis via modulation of Lin28B-let-7-HMGA2 signaling. Oncotarget *5*.

Yoon, J.-H., Abdelmohsen, K., Kim, J., Yang, X., Martindale, J.L., Tominaga-Yamanaka, K., White, E.J., Orjalo, A.V., Rinn, J.L., Kreft, S.G.*, et al.* (2013). Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. Nature Communications *4*, 2939.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. Science *318*, 1917-1920.

Yuan, J., Liu, X., Wang, T., Pan, W., Tao, Q., and cell …, Z.-W. (2017). The MBNL3 splicing factor promotes hepatocellular carcinoma by increasing PXN expression through the alternative splicing of lncRNA-PXN-AS1. Nature cell ….

Zou, Y., Chiu, H., Zinovyeva, A., Ambros, V., Chuang, F.C., and Chang, C. (2013). Developmental Decline in Neuronal Regeneration by the Progressive Change of Two Intrinsic Timers. Science *340*.