UCSF
UC San Francisco Previously Published Works

Title
Nuclear TARBP2 Drives Oncogenic Dysregulation of RNA Splicing and Decay

Permalink
https://escholarship.org/uc/item/2r90p5f9

Journal
Molecular Cell, 75(5)

ISSN
1097-2765

Authors
Fish, Lisa
Navickas, Albertas
Culbertson, Bruce
et al.

Publication Date
2019-09-01

DOI
10.1016/j.molcel.2019.06.001

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2020 September 05.

Published in final edited form as:

Mol Cell. 2019 September 05; 75(5): 967–981.e9. doi:10.1016/j.molcel.2019.06.001.

Nuclear TARBP2 drives oncogenic dysregulation of RNA splicing and decay

Lisa Fish[1,2,3,11], **Albertas Navickas**[1,2,3,11], **Bruce Culbertson**[1,2,3], **Yichen Xu**[2,3], **Hoang C.B. Nguyen**[5], **Steven Zhang**[1,2,3], **Myles Hochman**[1,2,3], **Ross Okimoto**[3,4], **Brian D. Dill**[6], **Henrik Molina**[6], **Hamed S. Najafabadi**[7,8], **Claudio Alarcón**[9,10], **Davide Ruggero**[2,3], **Hani Goodarzi**[1,2,3,*]

[1] Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA

[2] Department of Urology, University of California, San Francisco, San Francisco, CA 94158, USA

[3] Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA 94158, USA

[4]Department of Medicine, University of California, San Francisco, San Francisco, CA 94158, USA

[5] Laboratory of Systems Cancer Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

[6] Proteome Resource Center, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

[7] Department of Human Genetics, McGill University, Montreal, QC, Canada, H3A 0C7

[8] McGill University and Genome Quebec Innovation Centre, Montreal, QC, Canada, H3A 0G1

[9] Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520, USA

[10] Yale Cancer Biology Institute, Yale University, West Haven, CT 06516, USA

[11]These authors contributed equally

SUMMARY

Post-transcriptional regulation of RNA stability is a key step in gene expression control. We describe a regulatory program, mediated by the RNA binding protein TARBP2, that controls RNA stability in the nucleus. TARBP2 binding to pre-mRNAs results in increased intron retention,

[*] Lead contact: Hani Goodarzi, 600 16th St., San Francisco, CA 94158, Phone: 415-230-5189, hani.goodarzi@ucsf.edu.

AUTHOR CONTRIBUTIONS

Conceptualization, L.F. and H.G.; Methodology, L.F., A.N. and H.G; Formal Analysis, H.S.N. and H.G; Investigation L.F., A.N., B.C., Y.X., H.C.B.N., S.Z., M.H., R.O., B.D.D., H.M., C.A. and H.G.; Resources: D.R.; Writing - Original Draft, L.F. and H.G.; Writing - Review & Editing, L.F., A.N. and H.G.; Supervision, H.G.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

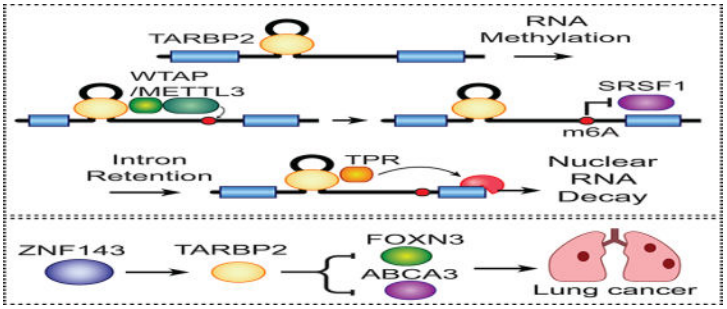
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

subsequently leading to targeted degradation of TARBP2-bound transcripts. This is mediated by TARBP2 recruitment of the m^{6}A RNA methylation machinery to its target transcripts, where deposition of m^{6}A marks influences the recruitment of splicing regulators, inhibiting efficient splicing. Interactions between TARBP2 and the nucleoprotein TPR then promote degradation of these TARBP2-bound transcripts by the nuclear exosome. Additionally, analysis of clinical gene expression datasets revealed a functional role for TARBP2 in lung cancer. Using xenograft mouse models, we find that TARBP2 impacts tumor growth in the lung, and that this is dependent on TARBP2-mediated destabilization of ABCA3 and FOXN3. Finally, we establish ZNF143 as an upstream regulator of TARBP2 expression.

eTOC blurb

Fish et al. show that TARBP2-mediated recruitment of the methyltransferase complex, and the subsequent deposition of m6A marks on TARBP2-bound transcripts, results in intron retention, leading to nuclear transcript decay. They also demonstrate that TARBP2 acts as a promoter of lung cancer growth through downregulation of ABCA3 and FOXN3 expression.

Graphical Abstract



INTRODUCTION

Post-transcriptional regulation of gene expression plays a major role in normal cell physiology and human diseases. The major molecular processes that are responsible for RNA turnover in the cytoplasm and the nucleus have been described in detail (Kilchert et al., 2016; Nasif et al., 2017). However, the regulatory programs that feed into these pathways to modulate transcript stability, and their collective role in shaping the cellular gene expression landscape, remain largely unexplored. Targeted intron retention has been described as one mechanism that modulates RNA degradation (Wong et al., 2016). In this pathway, transcripts with retained introns that are exported to the cytoplasm may be degraded by nonsense-mediated decay factors, or may be targeted by the nuclear RNA surveillance machinery prior to export. This latter mechanism has been reported for individual transcripts (Bergeron et al., 2015), as well as for controlling gene expression patterns during neuron development (Yap et al., 2012). However, the upstream regulatory programs that are involved in these nuclear RNA decay processes remain largely unknown. Here, we report the discovery and characterization of one such post-transcriptional regulatory network that functions in the nucleus to govern RNA stability.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We have previously described a regulatory pathway in which the double-stranded RNA binding protein TARBP2 binds and destabilizes its target transcripts through an unknown mechanism (Goodarzi et al., 2014). Here, we demonstrate that TARBP2 functions in the nucleus and modulates the stability of its regulon by influencing the rate of intron retention in its targets. Our findings reveal that nuclear TARBP2 recruits the RNA methylation machinery, resulting in local m^6A-mediated remodeling of splicing factors and impeding efficient processing of its target transcripts. RNA molecules with retained introns are then dispatched for degradation through interactions between TARBP2 and nuclear RNA decay factors. This regulatory program encompasses an interaction network between RNA modification, processing, and decay machineries, and reveals how they can function in concert to modulate the expression of a large regulon.

Our findings also highlight the emergence of RNA methylation as a major factor in post-transcriptional regulation of gene expression in the nucleus. The prevalent internal RNA modification mark N(6)-methyladenosine (m^6A) has been reported to play a role in regulating most facets of the RNA life cycle, including regulation of pre-mRNA splicing, mRNA stability, and mRNA translation (Lin et al., 2016; Liu et al., 2015; Wang et al., 2014a, 2015, 2014b; Xiao et al., 2016; Zhao et al., 2014). Work by us (Alarcón et al., 2015a, 2015b) and others (Ke et al., 2017; Knuckles et al., 2017) has established that m^6A marks are deposited in the nucleus and are proposed to function in many nuclear regulatory processes, including microRNA and messenger RNA processing. Despite the widespread use of these pathways, the underlying regulatory programs that influence m^6A deposition patterns across the transcriptome are poorly characterized. The TARBP2-mediated pathway described here adds a regulatory dimension to RNA methylation and its crucial role in targeted RNA turnover in the nucleus. Importantly, we have discovered that the increased activity of TARBP2 promotes lung cancer growth. Employing a network analytical approach, we have identified and functionally validated key factors that lie upstream and downstream of TARBP2 that take part in its oncogenic role in lung cancer. The importance of this TARBP2-mediated regulatory program in multiple cancer types further highlights its central role as a key regulator of gene expression.

RESULTS

TARBP2 binding results in increased intron retention and destabilization in the nucleus

To initially characterize the regulatory consequences of TARBP2 modulation we performed siRNA-mediated knockdown of TARBP2 followed by high-throughput RNA sequencing. We then asked whether TARBP2-bound transcripts, defined by analysis of TARBP2 HITS-CLIP data (Goodarzi et al., 2014), show a concerted change in abundance. Consistent with our previous findings obtained from microarrays, we observed that transcripts directly bound by TARBP2 were significantly upregulated when TARBP2 was silenced (Figure 1A). However, the molecular mechanisms linking TARBP2 binding to transcript destabilization were unknown.

We had observed that in our previously published TARBP2 HITS-CLIP data (Goodarzi et al., 2014) TARBP2 shows pervasive binding to intronic sequences (e.g. Figure S1A). This suggested that TARBP2 binds to pre-mRNAs and may therefore function in the nucleus by

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

influencing RNA processing and clearance. The major known pathway for RNA degradation in the nucleus involves the targeted destruction of incorrectly spliced transcripts by the RNA surveillance machinery (Kilchert et al., 2016). We hypothesized that TARBP2 may take advantage of this pathway by inhibiting efficient processing of introns to which it is bound, resulting in nuclear retention and degradation of its target transcripts. To assess the response of TARBP2-bound introns upon modulation of TARBP2 levels, we used high-throughput transcriptomic profiling measurements from control and TARBP2 knockdown cells to assess the changes in abundance of TARBP2 target transcripts at the exonic and intronic levels. We annotated TARBP2-bound introns and quantified their abundance relative to their flanking exons using a probabilistic model (MISO; (Katz et al., 2010)). To measure the global impact of TARBP2 silencing on the splicing of TARBP2-bound introns, we quantified the change in percent intron retention (PIR) across all TARBP2-bound introns in TARBP2 knockdown and control cells. As shown in Figure 1B, we observed a significant shift towards an increased rate of splicing of TARBP2-bound introns when TARBP2 is silenced.

Pervasive TARBP2 binding to intronic sequences implies that TARBP2 is present in the nucleus, which is further supported by a previous study that observed tagged TARBP2 in the nucleus of HeLa cells (Laraki et al., 2008). The Human Protein Atlas also provides immunofluorescence staining showing TARBP2 localized to the nucleoplasm of HeLa and MCF7 cells. To further verify that endogenous TARBP2 is also present in the nucleus of MDA-LM2 cells, we performed immunofluorescence staining followed by confocal microscopy to assess the cellular localization of TARBP2. Consistent with our model, TARBP2 was present in the nucleus as well as the cytoplasm of MDA-LM2 cells (Figure 1C). As will be discussed below, this observation was further confirmed through subcellular fractionation followed by label-free mass-spectrometry and western blotting for TARBP2.

The presence of TARBP2 in the nucleus, as well as its binding to intronic sequences, suggested that TARBP2 may mediate the processing and stability of its target transcripts in the nucleus. To investigate this possibility, we performed high-throughput sequencing on nuclear RNA from TARBP2 knockdown and control cells. We observed a highly significant increase in the expression of the TARBP2 regulon in the nucleus (Figure 1D); this effect was similar to but substantially stronger than the effect observed in total RNA (Figure 1A). To further investigate the effect of TARBP2 binding to introns we measured changes in intron retention by analyzing nuclear RNA-seq data from TARBP2 knockdown and control cells. Silencing TARBP2 resulted in a significant decrease in the abundance of TARBP2-bound introns compared to introns not bound by TARBP2 (Figures 1E, S1B). On average, TARBP2 target transcripts that have increased mature mRNA levels upon TARBP2 silencing showed a 5% reduction in retention of their TARBP2-bound introns ($P < 1^{-100}$). To test if the observed nuclear upregulation of the TARBP2 regulon is due to post-transcriptional modulation of RNA stability, we performed whole-genome nuclear transcript stability measurements by using α-amanitin to inhibit RNA polymerase II and gene expression profiling to assess changes in relative transcript stability in TARBP2 knockdown and control cells. Consistent with our previous observations (Goodarzi et al., 2014), we noted a significant enrichment of TARBP2-bound transcripts among those stabilized in the nucleus of TARBP2 knockdown cells, providing evidence that TARBP2 binding leads to transcript destabilization in the nucleus (Figure 1F).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Nuclear TARBP2 interacts with mRNA processing and export factors

In order to identify the molecular components through which nuclear TARBP2 causes its regulatory effects, we carried out an unbiased search for its interacting protein partners. We performed immunoprecipitation of both nuclear and cytoplasmic TARBP2, along with an IgG control, to identify proteins that interact with TARBP2 in the nucleus (Figure 2A). We searched for RNA-binding protein complexes that were significantly overrepresented in the TARBP2 immunoprecipitation samples compared to IgG co-precipitated proteins (StringDB; (Szklarczyk et al., 2015)). Ranking high on the list of statistically significant complexes were two involved in RNA processing: a complex containing the RNA processing factor WTAP, and another containing the nuclear pore-associated protein TPR (Figure S2A–B). This analysis suggested that TARBP2 may produce its effect on RNA stability through interactions with these proteins and their associated pathways.

TARBP2 recruits m^{6}A methylation machinery to mark target transcripts

To examine the role of WTAP in modulating expression of the TARBP2 regulon, we carried out siRNA-mediated knockdown of WTAP followed by RNA-seq to quantify both the expression of TARBP2 targets and the processing of their introns. As shown in Figure 2B, silencing WTAP resulted in a significant increase in the expression of transcripts bound by TARBP2. Importantly, these TARBP2-bound transcripts were significantly overrepresented in the set of transcripts upregulated in both TARBP2 and WTAP knockdown cells, and the gene expression changes resulting from TARBP2 and WTAP knockdown are well correlated ($R = 0.27$; Figure 2C). Moreover, the upregulation of the TARBP2 regulon upon WTAP silencing coincided with an increase in the splicing of TARBP2-bound introns (75% of introns with PIR below zero, $P < 1e\text{-}100$; Figure 2D). TARBP2 target transcripts with increased mature mRNA levels in WTAP knockdown cells had a 9% average reduction in PIR for their TARBP2-bound introns ($P < 1e\text{-}100$). Furthermore, analysis of a previously published WTAP PAR-CLIP dataset (Liu et al., 2014) revealed a significant overlap between TARBP2-bound introns and WTAP binding sites located in expressed introns and their flanking exons (Figure S2C). We included flanking exons of the TARBP2-bound introns in this analysis as many regulatory factors that influence intron splicing, including m6A methylation, occur in or interact with exonic regions. Together, these data show that WTAP silencing results in decreased intron retention and increased expression of the TARBP2 regulon, providing evidence that WTAP is a component of this TARBP2-mediated regulatory pathway.

The observed change in intron retention for TARBP2-bound introns in response to WTAP knockdown is consistent with the known function of WTAP as an RNA processing factor. However, WTAP also serves as the regulatory component of the m^{6}A methyltransferase complex (Liu et al., 2014; Ping et al., 2014), and the extent to which these two functions overlap is not known. Thus, the impact of WTAP on the TARBP2 regulon may be dependent on or independent of its role in RNA methylation. To address this question, we first asked whether the interaction between TARBP2 and WTAP has an impact on the methylation status of TARBP2-bound transcripts. We analyzed our previously published nuclear m^{6}A co-immunoprecipitation followed by sequencing data (MeRIP-seq; (Alarcón et al., 2015b)) and we observed a highly significant overlap between introns that are bound by TARBP2 and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

those that contain an m^{6}A mark (Figure 3A). While less than 10% of all expressed introns (and their flanking exons) show evidence of m^{6}A methylation, more than half of TARBP2-bound introns contain methylation marks (Figure 3A). These observations suggest a model where TARBP2-mediated recruitment of the methyltransferase complex results in the methylation of its target transcripts. In support of this model, we also observed that METTL3, the enzymatic component of the methyltransferase complex, co-immunoprecipitated with TARBP2, providing further evidence that TARBP2 interacts with the m^{6}A methyltransferase complex (Figure 3B). Furthermore, we observed a significant overlap between expressed introns (and their flanking exons) bound by TARBP2 and those bound by METTL3 (PAR-CLIP; (Liu et al., 2014)) (Figure 3C, example shown in Figure S3A).

In order to verify that the regulatory effects of WTAP are mediated through m^{6}A RNA methylation, and given that TARBP2 interacts with METTL3, we also analyzed high-throughput RNA-seq data from METTL3 knockdown cells. In this data we observed a significant increase in the expression of the TARBP2 regulon (Figure 3D). Consistent with this, we observed a significant decrease in retention of TARBP2-bound introns compared to introns not bound by TARBP2 in nuclear RNA from METTL3 knockdown cells (Figure 3E). To assess the direct effect of TARBP2 and METTL3 on intron splicing, we performed reporter assays. We generated MDA-LM2 cells with a stably integrated splicing reporter construct, and then used qRT-PCR to determine the relative level of spliced and unspliced reporter RNA in these cells with siRNA-mediated knockdown of TARBP2, METTL3 or control cells (see Methods). This assay showed a significant increase in spliced reporter transcript in TARBP2 and METTL3 knockdown cells compared to control cells (Figure S3B). Using this same assay, in TARBP2 and METTL3 knockdown cells we also observed a significant decrease in spliced reporter RNA in cells that had a deleted TARBP2 binding region in the reporter intron compared to the non-mutated reporter (Figure 3F). This indicates that decreased intron retention upon reduced levels of TARBP2 and METTL3 is contingent on the presence of TARBP2 binding sites. Finally, to test the causal link between TARBP2 binding and the methylation of its target RNAs, we performed MeRIP-Seq in TARBP2 knockdown and control cells. Consistent with a role for TARBP2 in recruiting the methyltransferase complex, in TARBP2 knockdown cells we observed a significant decrease in the m^{6}A signal in the TARBP2-bound introns relative to other expressed introns (Figure 3G).

TARBP2-dependent m^{6}A methylation impacts intron retention and RNA stability

It has been demonstrated that RNA methylation can occur co-transcriptionally, and that m^{6}A marks can be detected in chromatin-associated RNA (Ke et al., 2017). Analysis of these chromatin-associated RNA m^{6}A marks (CA-m^{6}A) revealed that there is a significant enrichment of TARBP2-bound introns among chromatin-associated methylated introns (Figure S3C). In contrast, CA-m^{6}A marks are largely absent in introns with no evidence of TARBP2 binding (background introns). This suggested that TARBP2 may regulate m^{6}A deposition co-transcriptionally. In support of this observation, we blotted for TARBP2 in different sub-cellular compartments, and detected TARBP2 in the chromatin-associated protein fraction, and, to a lesser extent, the soluble nuclear fraction (Figure 3H). This result

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

demonstrates that a portion of nuclear TARBP2 is associated with chromatin, which is consistent with a role for TARBP2 in promoting the m^6A methylation of nascent RNA.

To independently confirm the intron retention analysis of our RNA-seq data, we randomly selected a number of TARBP2-bound introns with known methylation sites (CA-m6A and MeRIP-Seq; (Alarcón et al., 2015b; Ke et al., 2017)). We then used qRT-PCR to measure TARBP2-dependent relative changes in retention of this set of introns using exon-exon and exon-intron spanning primers. In all but one of these cases, we observed a significant decrease in intron retention upon TARBP2 knockdown (Figure S3D). We also used qRT-PCR to confirm that these decreases in intron retention were accompanied by increases in the levels of these mature mRNAs (Figure S3E). As expected from our model, we also noted a significant anti-correlation between changes in intron retention and RNA expression for these transcripts (Figure S3F).

Taken together, our findings support a molecular mechanism of action where TARBP2-mediated methylation of introns interferes with splicing. Based on this model, it is plausible that pre-mRNA m^6A marks may modulate local binding of regulators of the splicing machinery. To address this possibility, we compiled a list of RNA-binding proteins that differentially bind methylated RNA (Alarcón et al., 2015a; Edupuganti et al., 2017; Liu et al., 2015). We then systematically analyzed the distribution of binding sites of each RBP (derived from CLIP-seq data and known RBP binding motifs, 33 RBPs were included in this analysis) to determine if each candidate RBP binds to introns (and their flanking exons) that are also bound by TARBP2 and/or contain m6A marks. We note that this analysis was limited to proteins with available CLIP-seq data, and therefore does not include all m^6A binding proteins. This analysis identified SRSF1 (Serine and arginine rich splicing factor 1), a major modulator of both pre-mRNA splicing and alternative splicing (Das and Krainer, 2014), and a factor that exhibits decreased binding to methylated RNA (Edupuganti et al., 2017), as a top candidate. To determine if SRSF1 and TARBP2 bind the same transcripts we performed CLIP for SRSF1 in MDA-LM2 breast cancer cells and H1299 lung cancer cells, and found a significant overlap in SRSF1 and TARBP2-bound introns (and flanking exons) as well as SRSF1-bound and m^6A mark-containing introns (and flanking exons) (Figure 3I). These analyses suggest a model where SRSF1-dependent intron splicing is inhibited by TARBP2-dependent m^6A methylation of introns. To evaluate this hypothesis, we performed RNA-seq on SRSF1 knockdown and control cells, and, consistent with our model, we observed a significant decrease in the expression of the TARBP2 regulon upon SRSF1 knockdown (Figure S3G). Moreover, our analysis shows that transcripts bound by TARBP2 were significantly overrepresented in the set of transcripts that was both upregulated in TARBP2 knockdown cells and downregulated in SRSF1 knockdown cells (Figure 3J). Based on our analyses, in addition to SRSF1, other regulators of RNA splicing that are known to differentially bind methylated RNA may also contribute to TARBP2-dependent intron retention. One potential additional factor is the splicing regulator HNRNPC, which has been shown to preferentially bind m^6A methylated RNA (Liu et al., 2015). We observed a significant overlap between the target introns bound by HNRNPC (Zarnack et al., 2013) and those bound by TARBP2 (Figure S3H). Consistently, we also observed an increase in the expression of the TARBP2 regulon when HNRNPC is silenced (Liu et al., 2015) (Figure S3I). We similarly assessed the role of YTHDC1, a known nuclear m^6A reader that has been

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

implicated in splicing regulation (Xiao et al., 2016; Xu et al., 2014). However, we observed only a slight overlap between TARBP2 and YTHDC1 binding on introns (Xiao et al., 2016). Consistently, YTHDC1 knockdown resulted in little change in the expression of the TARBP2 regulon (data not shown). Together, our findings implicate SRSF1 as a key splicing regulator that is repelled upon methylation of its binding sites on TARBP2-bound introns. Subsequently, decreased SRSF1 binding results in increased intron retention and decreased expression of the mature transcript. In addition to SRSF1, HNRNPC may play a minor role in this regulatory process.

TARBP2 delivers its target transcripts to the nuclear RNA surveillance complex for degradation

In our TARBP co-immunoprecipitation data, TPR, a nuclear pore-associated factor, was the highest-ranking TARBP2 interacting protein in the nucleus (Figure S2A). Given the known role of this protein in nuclear RNA surveillance and degradation of mis-spliced transcripts (Coyle et al., 2011; Krull et al., 2004; Rajanala and Nandicoori, 2012), its interaction with TARBP2 suggested a direct mechanism for the nuclear retention and degradation of TARBP2-bound transcripts. Consistent with this, the TARBP2 regulon was significantly enriched among transcripts that were upregulated upon TPR knockdown (Figure 4A). Gene expression changes in TARBP2 and TPR knockdown cells were also positively correlated, highlighting the functional overlap of these two proteins (Figure 4B). Furthermore, TARBP2-bound transcripts were significantly enriched among the set of transcripts that was upregulated in both TARBP2 and TPR knockdown cells (Figure 4B). If these regulatory consequences of TPR are mediated through its function as a component of the nuclear RNA surveillance machinery, other factors in this complex should have a similar impact on the TARBP2 regulon. To evaluate the role of the nuclear RNA surveillance machinery in TARBP2-mediated transcript destabilization, we analyzed previously reported iCLIP data for EXOSC10 (Macias et al., 2015), a catalytic component of the nuclear exosome complex. As shown in Figure S4A, TARBP2-bound transcripts are significantly enriched among those that are also bound by EXOSC10, consistent with these transcripts being targeted by the surveillance machinery for degradation. Moreover, an additional independent RNA-seq dataset from cells with EXOSC10 knockdown (Macias et al., 2015) revealed a significant increase in the expression of the TARBP2 regulon (Figure S4B). We confirmed this observation by performing RNA-seq in cells with knockdown of EXOSC10, and also tested the effect of knockdown of XRN2 (a nuclear 5’ to 3’ exonuclease (Miki and Großhans, 2013)). This revealed that knockdown of EXOSC10 or XRN2 resulted in a significant increase in the expression of the TARBP2 regulon (Figures 4C, S4C). To further investigate the effect of an additional nuclear RNA decay factor and a cytoplasmic RNA decay factor on the TARBP2 regulon, we knocked down RBM7 (a component of the NEXT complex (Meola et al., 2016)), as well as the cytoplasmic RNA decay factor UPF1 (Figure S4D). We noted a significant increase in the expression of the TARBP2 regulon in cells with RBM7 knockdown, and no significant difference in the expression of the TARBP2 regulon in cells with UPF1 knockdown. These results are consistent with the model that destabilization of the TARBP2 regulon is contingent on the catalytic activity of the nuclear RNA decay machinery.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Upregulation of the TARBP2 pathway is associated with lung cancer

We have previously described a role for aberrant TARBP2 activity in metastatic breast cancer (Goodarzi et al., 2014). Here, we sought to identify the broader role of this pathway in normal cell physiology and human disease. To address this problem, we performed an unbiased search for evidence of aberrant TARBP2 activity across the human cancers profiled in The Cancer Genome Atlas (TCGA). Consistent with our previous findings, we observed a significant association between the TARBP2 gene expression signature and breast cancer. However, the TARBP2 signature also showed broad upregulation in several other cancer types, with the strongest association observed in lung cancer (Figure S5A). This observation was also validated in independent lung cancer datasets, which showed strong upregulation of TARBP2 (Figures 5A, S5B). Importantly, we also observed a strong association between TARBP2 expression and survival in lung cancer patients (Figure 5B). To independently confirm this observation, we used qRT-PCR to measure TARBP2 mRNA levels in lung tumor samples from a cohort of lung cancer patients (including 20 stage I, 14 stage II, and 6 stage III) as well as lung tissue from healthy individuals. As shown in Figure 5C, we observed a substantial and significant upregulation of TARBP2 in lung adenocarcinoma compared to healthy tissue samples.

Consistent with our observations in breast cancer cell lines, silencing TARBP2 in H1299 lung cancer cells resulted in both a significant upregulation and stabilization of the TARBP2 regulon (Figure 5D–E). Similarly, in the lung cancer cell lines A549 and H1650, we observed an upregulation of the TARBP2 regulon upon TARBP2 knockdown (Figure S5C–D). Together, these analyses provide evidence that TARBP2 is strongly associated with lung cancer and that TARBP2-mediated modulation of RNA stability occurs in lung cancer cells.

TARBP2 promotes lung cancer in *in vivo* models

As our analysis of clinical datasets provides evidence that TARBP2 plays a role in lung cancer, we sought to experimentally test this hypothesis using xenograft mouse models. Initially, we injected TARBP2 knockdown and control H1299 lung cancer cells into the venous circulation of immunodeficient mice and then measured cancer cell growth in the lung over time using *in vivo* bioluminescence imaging. While TARBP2 knockdown resulted in only a modest decrease in *in vitro* cell proliferation (Figure S5E), we observed a significant reduction in growth in the lung by cells with TARBP2 knockdown compared to control cells (Figure 5F). Next, we used an orthotopic xenograft assay to directly assess cancer cell growth in the lung. This revealed a significant decrease in tumor growth of H1299 TARBP2 knockdown cells compared to control cells (Figure S5F). We also performed this assay using an independent lung cancer cell line, H1975, and similarly observed a significant decrease in cancer cell growth in the lung of TARBP2 knockdown compared to control cells (Figure 5G), with no significant difference observed in their *in vitro* proliferation rates (Figure S5E). Consistent with these results, stable overexpression of TARBP2 in H1975 cells resulted in a significant increase in cancer cell growth in the lung compared to control mCherry overexpressing cells (Figure 5H), with no significant difference observed in their *in vitro* proliferation rates (Figure S5E). Together, these results provide evidence that TARBP2 acts as a promoter of lung cancer cell growth in the lung *in vivo.*

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In order to identify TARBP2 targets that act downstream of TARBP2 to impact lung cancer, we searched for transcripts that were directly bound by TARBP2, had TARBP2-dependent decreased expression and stability, and were negatively correlated with TARBP2 expression in clinical lung cancer datasets (Figure S6A). From this list, we selected the four highest ranking targets that also showed evidence of methylation and TARBP2-dependent intron retention, namely FOXN3, ABCA3, DAB2, and STK10. We tested the impact of silencing these genes on lung cancer growth in xenograft models by injecting H1299 cells stably expressing an shRNA against each of these genes. As shown in Figure 6A, silencing FOXN3 and, to a lesser extent, silencing ABCA3, significantly increased lung cancer growth, while decreasing STK10 and DAB2 expression had no significant effect. Moreover, this effect was independent of *in vitro* cell proliferation rates, which showed no significant change upon target gene knockdown (Figure S6B). We also performed qRT-PCR to measure the relative levels of the mature and pre-mRNA of these targets in TARBP2 knockdown and control H1299 cells, and observed a significant increase in the mature mRNA levels and a decrease in the relative pre-mRNA levels of ABCA3 and FOXN3 upon TARBP2 knockdown (Figure S6C). To further assess the clinical relevance of these functional target genes in human disease, we performed a set of additional analyses using clinical datasets. We analyzed a dataset of gene expression profiles from a large cohort of matched normal and lung tumor samples collected from patients (Kim et al., 2013), and we confirmed that TARBP2 is also significantly upregulated in this data (Figure 6B). We then assessed the changes in the expression of FOXN3 and ABCA3 in this same dataset, and consistent with their proposed roles as tumor suppressors, we observed a highly significant reduction in their expression in lung cancers, and found that their expression was significantly correlated with that of TARBP2 (Figure S6D). To test if FOXN3 acts downstream of TARBP2 in promoting lung cancer growth, we performed *in vivo* orthotopic xenograft assays using H1299 lung cancer cells with knockdown of FOXN3 and TARBP2 or knockdown of TARBP2 alone, along with control cells. Consistent with our previous results, this experiment showed a significant decrease in growth in the lung of cells with TARBP2 knockdown only, and an increase in growth in the lung of cells with knockdown of both FOXN3 and TARBP2 compared to control cells (Figure 6C). Together, these data are consistent with a model where TARBP2 decreases expression of FOXN3 and ABCA3, leading to increased lung cancer growth.

The transcription factor ZNF143 drives the aberrant upregulation of TARBP2

While our findings provide a molecular understanding of how TARBP2 promotes cancer growth and progression, they did not explain how cancer cells achieve TARBP2 overexpression. Analysis of the TCGA-LUAD (lung adenocarcinoma) dataset showed a general association between TARBP2 expression and TARBP2 genomic copy number (Figure S7A). However, copy number variation alone is not sufficient to explain the magnitude of TARBP2 upregulation in lung cancer. In order to identify the regulatory pathway that drives TARBP2 overexpression in lung cancer, we performed a systematic search for known transcription factors that were significantly co-expressed with TARBP2 in multiple independent cancer datasets. We also performed promoter sequence analysis and ChIP-seq data mining (ENCODE) to identify transcription factors that potentially act as upstream regulators of TARBP2. This exercise yielded a list of five potential candidates. To ask if any of these candidates regulate TARBP2 expression, we performed siRNA-mediated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

knockdown of each gene in H1299 lung cancer cells followed by qRT-PCR (Figure S7B). Of these candidates, only ZNF143 knockdown resulted in a significant reduction in TARBP2 expression, consistent with ZNF143 controlling TARBP2 transcription. This result was observed with independent ZNF143-targeting siRNAs in additional lung and breast cancer cell lines (Figures 7A, S7C). Moreover, ChIP-seq datasets from multiple cell lines show evidence of strong binding of ZNF143 at the TARBP2 promoter, with the ChIP peak in this region containing a close match to the ZNF143 consensus binding site (Figure 7B). Our analysis also revealed a significant correlation between expression of ZNF143 and TARBP2 in lung cancer gene expression data (Figure S7D). Finally, consistent with ZNF143 playing a role in modulating TARBP2 transcription, we found a significant association between ZNF143 expression and survival in lung cancer patients, as well as a significant upregulation of ZNF143 levels in lung cancers compared to matched normal tissue in the TCGA-LUAD (lung adenocarcinoma) dataset (Figures 7C–D). Together, these results provide strong evidence that the transcription factor ZNF143 increases the expression of TARBP2 in breast and lung cancers.

DISCUSSION

Here, we describe an oncogenic post-transcriptional regulatory program controlled by the double stranded RNA binding protein TARBP2. First described as a protein that binds the HIV TAR element, TARBP2 also has a role in miRNA processing (Chendrimada et al., 2005; Gatignol et al., 1991; Kim et al., 2014). Our previous findings demonstrated that TARBP2 regulates RNA stability through the direct binding of RNA structural elements on hundreds of transcripts (Goodarzi et al., 2014). In this study, we dissect the molecular mechanisms through which TARBP2 controls transcript stability, and show that TARBP2 directly controls the stability of its bound targets via co-transcriptional recruitment of the METTL3 methyltransferase complex, resulting in intron methylation and subsequent retention of the intron followed by degradation of the transcript by the nuclear exosome.

Intron retention is a well-documented mechanism of regulating RNA stability. Here, we establish a link between TARBP2 intronic binding, m^6A methylation, and controlled intron retention leading to transcript degradation in the nucleus. Although RNA methylation marks impact a variety of developmental and disease processes (Zhang et al., 2017a, 2017b; Zhao et al., 2014), their mechanistic effects have not been fully explored. One known mechanism linking m^6A methylation of transcripts to their stability is a cytoplasmic process mediated by the m^6A binding protein YTHDF2, in which m^6A-containing transcripts are bound by YTHDF2 in the cytoplasm, which then recruits the CCR4-NOT complex to accelerate their degradation (Du et al., 2016). The pathway we describe here is distinct in that it occurs in the nucleus, implying that it may act on different sets of transcripts. Interestingly, in S. *pombe,* Mmi1, a YTH domain-containing protein binds specific introns, resulting in their retention and subsequent targeted nuclear degradation by the exosome. Although methylation of the target introns has not been demonstrated, this suggests existence of a more general link between intron methylation, retention, and nuclear decay (Kilchert et al., 2015). Building on this mechanism, we also found that a fraction of TARBP2 is associated with chromatin, and therefore it is plausible that TARBP2 promotes intron methylation co-transcriptionally, consistent with a report that m^6A marks are deposited on nascent pre-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

mRNA(Ke et al., 2017). We speculate that the mode of TARBP2-dependent post-transcriptional regulation we describe here may be advantageous because it allows for the fast decoupling of expression of the TARBP2 regulon from transcriptionally controlled levels. This may be beneficial for promoting oncogenesis and metastasis as it could allow for rapid adaptation to changes in the microenvironment.

Although we found a highly significant association between TARBP2 intron binding and RNA methylation, the methylation sites do not necessarily overlap with TARBP2 binding sites. This suggested that additional factors interact with the methylated sites to promote intron retention. By analyzing our data, along with publically available datasets, we found strong associations between TARBP2-mediated intron retention and the splicing regulator proteins SRSF1 and HNRNPC. Our results are consistent with a mechanism where methylation interferes with the ability of SRSF1 to bind and promote intron processing, and potentially enhances the ability of HNRNPC to bind and inhibit intron processing. HNRNPC has been shown to compete with U2AF65 (Zarnack et al., 2013), a function that could inhibit spliceosome assembly, leading to local intron retention.

Transcripts with retained introns may be excluded from cytoplasmic export by TPR, a protein component of the nuclear basket that has a known role in impeding the export of intron containing RNAs (Coyle et al., 2011; Rajanala and Nandicoori, 2012). In this study, we found a physical interaction between TARBP2 and TPR. This TARBP2-TPR interaction suggests that TARBP2 is in close proximity to the nuclear basket, and therefore it is possible that TARBP2-bound intron-containing transcripts are immediately blocked from cytoplasmic export by proximity to TPR. We also identified EXOSC10 and XRN2, catalytic factors of the canonical nuclear decay machinery, as factors responsible for the nuclear degradation of the TARBP2 regulon. Additionally, we found that RBM7 plays a role in this process. RBM7 is part of the NEXT complex, acting as an adapter for the nuclear exosome, and has also been found to associate with SF3b, a component of the spliceosome, and it has been proposed that this association recruits unspliced RNAs to the nuclear exosome (Falk et al., 2016).

Furthermore, we provide evidence for a functional role for a TARBP2-mediated RNA decay pathway in lung cancer. Intriguingly, we had previously observed that TARBP2 promotes metastatic colonization of the lung by breast cancer cells. Here, we find that the TARBP2 signature is enriched in clinical lung cancer gene expression datasets, and we demonstrate that TARBP2 enhances lung cancer growth *in vivo,* suggesting that the gene expression pattern controlled by TARBP2 is highly suited for promoting oncogenic growth in the lung microenvironment. Also consistent with our results, a previous study reported that knockdown of TARBP2 in H1299 lung cancer cells reduced cell invasion and migration (Shi et al., 2016). Furthermore, our analyses show a robust association between TARBP2 expression and clinical outcome in lung cancer. These findings are consistent with the view that gene expression programs that promote primary tumor growth may also be critical in promoting metastatic colonization of that same organ.

We have also found that in lung cancer, TARBP2 downregulates ABCA3 and FOXN3 expression, leading to increased cancer growth in the lung. ABCA3 is an ATP binding

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cassette lipid transport protein that is necessary for normal secretion of lung surfactant (Shulenin et al., 2004). Deletion of ABCA3 in mouse genetic models leads to lung tissue injury, inflammation and subsequent proliferation of new cells from progenitors (Rindler et al., 2017). It is possible that these inflammatory and proliferative processes could promote lung cancer growth. It is also possible that ABCA3 directly suppresses cancer cell growth through modulating lipid transport. FOXN3, a second functional TARBP2 target, is a transcriptional repressor that also acts as a cell cycle checkpoint regulator (Pati et al., 1997; Scott and Plon, 2005). Therefore, it is plausible that knockdown of FOXN3 promotes uncontrolled cell division leading to oncogenesis. Although we did not observe a significant change in the *in vitro* proliferation rate of FOXN3 knockdown cells, it is possible that lung-specific microenvironmental cues are required for this effect. Consistent with our findings in lung cancer, downregulation of FOXN3 has been reported to promote proliferation of liver and colon cancer cells (Dai et al., 2016; Sun et al., 2016). Intriguingly, it has been reported that a G to A single-nucleotide polymorphism in the first intron of FOXN3 results in higher FOXN3 expression—it is conceivable that this is a result of disruption of the TARBP2-mediated pathway (Karanth et al., 2016). While ABCA3 and FOXN3 transcripts are directly bound by TARBP2, the possibility remains that their expression is modulated indirectly by TARBP2.

Finally, we have identified ZNF143 as an upstream regulator of TARBP2 expression. Consistent with this, a previous study found an association between high ZNF143 protein levels and poor survival in lung adenocarcinoma (Kawatsu et al., 2014). Intriguingly, a study has identified small molecules that inhibit ZNF143 activity (Haibara et al., 2017), pointing towards a possible avenue for inhibition of TARBP2 pro-oncogenic activity.

Taken together, our study describes a post-transcriptional regulatory program that establishes a functional link between the RNA methylation machinery, regulators of RNA splicing, and components of the nuclear RNA surveillance complex. Together, these processes combine to create a regulatory mechanism, orchestrated by the RNA-binding protein TARBP2, that modulates the expression of a large set of transcripts in the nucleus. Linking TARBP2 to both lung and breast cancer progression emphasizes its importance in shaping the gene expression landscape of the cell. However, the same mechanisms might also be employed by other post-transcriptional regulators to modulate expression of their associated regulons. As such, a broader understanding of controlled intron retention and its underlying molecular mechanisms is a crucial step towards achieving a more detailed view of post-transcriptional regulation, as well as exposing new vulnerabilities that can exploited to counter human disease.

STAR Methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Hani Goodarzi (hani.goodarzi@ucsf.edu).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Experimental Model and Subject Details

Animals—Seven- to twelve-week-old age-matched NOD/SCID gamma mice (Jackson Labs) were used for lung colonization assays. Six- to eight-week-old NOD/SCID gamma mice (Jackson Labs) were used for lung xenograft experiments. Male mice were used for assays with H1299 cells, and female mice were used for assays with H1975 cells. Mice were randomly assigned to experimental cohorts. Specific-pathogen-free conditions and facilities were approved by the American Association for Accreditation of Laboratory Animal Care. Surgical procedures were reviewed and approved by the UCSF Institutional Animal Care and Use Committee (IACUC), protocol #AN107889–03 and protocol #AN179718.

Cell lines—All cells were cultured in a 37°C 5% CO2 humidified incubator. The MDA-MB-231 breast cancer cell line, its highly metastatic derivative, MDA-LM2 (Minn et al., 2005), and 293LTV cells were cultured in DMEM medium supplemented with 10% FBS, glucose (4.5g/L), L-glutamine (4mM), sodium pyruvate (1mM), penicillin (100 units/mL), streptomycin (100 μg/mL) and amphotericin (1μg/mL) (Gibco). The H1650 and H1299 lung cancer cell lines were cultured in RPMI-1640 medium supplemented with 10% FBS, glucose (2g/L), L-glutamine (2mM), sodium pyruvate (1mM), penicillin (100 units/mL), streptomycin (100 μg/mL) and amphotericin (1μg/mL) (Gibco). The A549 lung cancer cell line was cultured in F-12K medium supplemented with 10% FBS, penicillin (100 units/mL), streptomycin (100 μg/mL) and amphotericin (1μg/mL) (Gibco). The MDA-MB-231, MDA-LM2, and H1975 cells lines are female. The H1650, H1299, and A549 cell lines are male. All cell lines were routinely screened for mycoplasma with a PCR-based assay.

Method Details

Cell culture—All cells were cultured in a 37°C 5% CO2 humidified incubator. The MDA-MB-231 breast cancer cell line, its highly metastatic derivative, MDA-LM2 (Minn et al., 2005), and 293LTV cells were cultured in DMEM medium supplemented with 10% FBS, glucose (4.5g/L), L-glutamine (4mM), sodium pyruvate (1mM), penicillin (100 units/mL), streptomycin (100 μg/mL) and amphotericin (1μg/mL) (Gibco). The H1299, H1975, and H1650 lung cancer cell lines were cultured in RPMI-1640 medium supplemented with 10% FBS, glucose (2g/L), L-glutamine (2mM), sodium pyruvate (1mM), penicillin (100 units/mL), streptomycin (100 μg/mL) and amphotericin (1μg/mL) (Gibco). The A549 lung cancer cell line was cultured in F-12K medium supplemented with 10% FBS, penicillin (100 units/mL), streptomycin (100 μg/mL) and amphotericin (1μg/mL) (Gibco).

shRNA and siRNA-mediated knockdown—For stable knockdown of target genes, shRNAs targeting the genes of interest were cloned into the EcoRI and AgeI sites of the pLKO.1 vector (Addgene plasmid #10878; Moffat et al., 2006), see Table S2 for shRNA sequences. The shRNA constructs were then packaged using the ViraSafe lentiviral packaging system (Cell Biolabs, Inc.) by using lipofectamine 2000 (Invitrogen) and Opti-MEM (Invitrogen) to transfect the shRNA constructs along with the packaging plasmids into 293LTV cells (Cell Biolabs, Inc.). Virus was harvested 48 hours post-transfection and passed through a 0.45um filter. Target cells constitutively expressing luciferase were then transduced for 6–8 hours with the filtered virus in the presence of 8ug/mL polybrene (Millipore). Transduced cells were selected by treatment with 1.5ug/mL puromycin for 2–3

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

days. For the double knockdown cells, shRNAs were cloned into pLKO.1 with a blasticidin selection marker (Addgene plasmid #26655; Bryant et al., 2010) and the resulting virus was transduced into shTARBP2 cells and then selected with 10ug/mL blasticidin for 5–7 days. Three independent TARBP2-targeting shRNAs were used to knock down TARBP2 in the different cell lines described in this study. For MDA-LM2, shTARBP2–2 was used, for H1299, shTARBP2–1 was used, and for H1975, shTARBP2–3 was used. Knockdown of target genes was assessed by qRT-PCR as described below.

For transient knockdown of target genes, siRNAs (IDT DNA; see Table S3) were used. 100pmol of each siRNA was transfected into 1×10^5 cancer cells using lipofectamine 2000 and Opti-MEM (Invitrogen) per the manufacturer's protocol. Cells were harvested 48–72 hours post-transfection. Knockdown of target genes was assessed by qRT-PCR as described below.

RNA Isolation—Total RNA for RNA-seq and quantitative RT-PCR assays was isolated using the Norgen Biotek total RNA isolation kit with on-column DNase treatment per the manufacturer's protocol.

Quantitative RT-PCR—Transcript levels were measured using quantitative RT-PCR by reverse transcribing total RNA to cDNA (Superscript III, Invitrogen), then using fast SYBR green master mix (Applied Biosystems) or Perfecta SYBR green supermix (QuantaBio) per the manufacturer's instructions. HPRT1 and 18S were used as endogenous controls.

Metastatic colonization assays—Seven- to twelve-week-old age-matched NOD/SCID gamma mice (Jackson Labs) were used for lung colonization assays. Male mice were used for assays with H1299 cells, and female mice were used for assays with H1975 cells. In all cases, 5×10^4 cancer cells constitutively expressing luciferase were suspended in 100 μL PBS and then injected via tail-vein. Each cohort contained 4–5 mice. Cancer cell growth was monitored *in vivo* at the indicated times by retro-orbital injection of 100ul of 15mg/mL luciferin (Perkin Elmer) dissolved in 1X PBS, and then measuring the resulting bioluminescence with an IVIS instrument and Living Image software (Perkin Elmer).

Orthotopic lung xenograft assays—Six- to eight-week-old NOD/SCID gamma mice (Jackson Labs) were used for lung xenograft experiments. Male mice were used for assays with H1299 cells, and female mice were used for assays with H1975 cells. To prepare cell suspensions for thoracic injection, we briefly trypsinized adherent cancer cells constitutively expressing luciferase, quenched them with 10% FBS RPMI media, and resuspended them in 1X PBS. Cells were pelleted again and mixed with Matrigel matrix (Corning) on ice to achieve a final concentration of 0.75×10^5 cells/μl. The Matrigel cell suspension was transferred into a 1-ml syringe and remained on ice until the time of implantation.

For orthotopic injection, mice were placed in the right lateral decubitus position and anesthetized with 2.5% inhaled isoflurane. A 1-cm surgical incision was made along the posterior medial line of the left thorax. Fascia and adipose tissue layers were dissected and retracted to expose the lateral ribs, the intercostal space, and the left lung parenchyma. Upon recognition of left lung respiratory variation, a 30-gauge hypodermic needle was used to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

advance through the intercostal space ~3 mm into the lung tissue. For human cancer cell lines, care was taken to inject 10μl (7.5×10^5 cells) of cell suspension directly into the left lung. The needle was rapidly withdrawn, and mice were observed for pneumothorax. Visorb 4/0 polyglycolic acid sutures were used for primary wound closure of the fascia and skin layer. Mice were observed post-procedure for 1–2 hours, and their body weights and wound healing were monitored weekly. Cancer cell growth was monitored *in vivo* at the indicated times by retro-orbital injection of 100ul of 15mg/mL luciferin (Perkin Elmer) dissolved in 1X PBS, and then measuring the resulting bioluminescence with an IVIS instrument and Living Image software (Perkin Elmer).

Histology—For gross macroscopic metastatic nodule visualization, mice lungs (from each cohort) were extracted at the endpoint of each experiment, and 5μm thick lung tissue sections were hematoxylin and eosin (H&E) stained. The number of macroscopic nodules was then recorded for each section. Unpaired t-test was used to test for significant variations.

Cancer cell proliferation—5×10^4 cells were seeded into three 6-well wells and subsequently were trypsinized and stained with trypan blue to determine cell viability. Viable cells were counted using a hemocytometer at day 1, day 3 and day 5. An exponential model was then used to fit a growth rate for each sample $(ln(N_{t-1}/N_1)= rt$ where t is measured in days). The experiment was performed in biological quadruplicates and an unpaired t-test was used to test for significant variations.

Immunofluorescence—MDA-LM2 cells were seeded and incubated for 48 hours in chamber slides, then fixed with 4% paraformaldehyde, washed 3x with PBST (0.1% Tween-20), and blocked for one hour in blocking buffer (5% goat serum, 0.2% fish skin gelatin, 0.2% tween-20). TARBP2 primary antibody (Proteintech 15753) was diluted 1:50 in blocking buffer and incubated with the cells at 4°C overnight. Cells were washed three times with P BST before incubation with anti-rabbit Cy3 secondary (1:1000) (Jackson) in blocking buffer at 37°C for one hour. Cells were washed three times with PBST, the second wash containing 2.5 μg/mL DAPI. Slides were mounted with ProLong Gold antifade reagent (Life Technologies), and images acquired with a Nikon Ti spinning disk confocal microscope at the UCSF Nikon Imaging Center.

Stability measurements—MDA-LM2 TARBP2 knockdown and control cells were treated with 10μg/mL α-amanitin (final concentration in the medium). After 9 hours, nuclear RNA was harvested from the cells using the Norgen Cytoplasmic and Nuclear RNA Purification Kit per the manufacturer's protocol. This RNA was prepared for microarray using TargetAmp-Nano Labeling Kit for Illumina (Epicentre). Labeled RNA was purified using RNeasy Minelute Kit (Qiagen) and submitted for analysis to the Rockefeller University genomics core facility using Illumina HT-12 v4 Expression BeadChip microarrays. The Lumi package in R was used to transform and normalize signal intensities.

Co-immunoprecipitation and mass spectrometry—MDA-LM2 cells (in biological replicates) were collected by scraping, then centrifuged to pellet. The pellets were then resuspended in buffer LB1 (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% Triton X-100 and 1X protease inhibitors), incubated on ice for 15 minutes,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and spun for 10 minutes at 400 × g at 4°C. The supernatant was collected (cytosolic fraction) and the nuclei were resuspended in M-PER (Thermo Scientific) plus protease inhibitors, incubated on ice for 10 minutes and spun at 20,000 × g for 5 minutes at 4°C. Immunoprecipitation of TARBP2 was carried out by first removing glycerol from anti-TARBP2 and rabbit IgG using an antibody purification kit (Abcam) and then conjugating the antibodies to epoxy magnetic beads using the dynabeads antibody coupling kit (Invitrogen), all according to the manufacturer's protocol. The cytosolic and nuclear lysates were then incubated with antibody-conjugated beads for three hours at 4°C wi th end-over-end rotation. The beads were then washed three times with ice-cold PBS supplemented with 150mM NaCl and then three times with ice-cold PBS.

Following immunoprecipitation, proteins were eluted from the antibody-bead conjugate by denaturation in 50μL 8M urea/0.1M ammonium bicarbonate for 30 minutes. Supernatant was removed following reduction (10 mM DTT) and alkylation (40 mM iodoacetamide). Proteins were digested with Endoproteinase LysC (Wako Chemicals) after dilution to 4M urea followed by trypsination (Promega) in 2M urea. The digestion was quenched with 5% formic acid (final concentration) and resulting peptide mixtures were desalted using in-house made C18 Empore (3M) StAGE tips (Rappsilber et al., 2007). Samples were dried and resolubilized in 2% acetonitrile and 2% formic acid and analyzed by reversed phase nano-LC-MS/MS (Ultimate 3000 coupled to a QExactive Plus, Thermo Scientific). After loading on a C18 PepMap trap column (5 μm particles, 100μm × 2 cm, Thermo Scientific) at a flow rate of 3 μl/min, peptides were separated using a 12 cm × 75μm C18 column (3 μm particles, Nikkyo Technos Co., Ltd. Japan) at a flow rate of 200 nL/min, with a gradient increasing from 5% Buffer B (0.1% formic acid in acetonitrile) / 95% Buffer A (0.1% formic acid) to 40% Buffer B / 60% Buffer A, over 75 minutes. All LC-MS/MS experiments were performed in data dependent mode with lock mass of m/z 445.12003. Precursor mass spectra were recorded in a 300–1400 m/z range at 70,000 resolution, and fragment ions at 17,500 resolution (lowest mass: m/z 100) in profile mode. Up to twenty precursors per cycle were selected for fragmentation and dynamic exclusion was set to 60 seconds. Normalized collision energy was set to 27.

Data were searched against a Uniprot human database (July 2014) using MaxQuant (version 1.5.0.30) software and Andromeda search engine (Cox et al., 2014). Oxidation of methionine and N-terminal protein acetylation were allowed as a variable, and cysteine carbamidomethylation as a fixed modification. An ion mass tolerance was set at 4.5 ppm for precursor and 20 ppm for fragment ions. Two missed cleavages were allowed for specific tryptic search. The "match between runs" option was enabled. False discovery rates for proteins and peptides were set to 1%. Protein abundances were represented by LFQ (label free quantitation) values. Data were filtered to exclude contaminants, and reverse database hits. LFQ values were log2(x) transformed and further used for t-test (Tyanova et al., 2016).

Co-immunoprecipitation of METTL3—Co-immunoprecipitation of TARBP2 and METTL3 was performed using lysate prepared from MDA-LM2 cells. Cells were washed with ice-cold 1X PBS, collected by scraping, then centrifuged to pellet. Cell pellet was resuspended in M-PER (Thermo Scientific) plus 1X protease inhibitors (Thermo Scientific) and incubated on ice 10 minutes. Lysate was clarified by centrifugation for 15 minutes at

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

20,000 × g at 4°C. TARBP2- or IgG-conjugated epoxy dynabeads (prepared as described above for co-immunoprecipitation and mass spec) were added to clarified lysate and incubated for two hours at 4°C with end-over-end rotation. Beads were then washed once with PBS supplemented with 150mM NaCl, then washed three times with 1X PBS. Proteins were eluted by resuspending beads in loading buffer (1x NuPAGE LDS loading buffer, 50mM DTT) and incubating for 10 minutes at 70°C. The presence of METTL3 in input, flowthrough and eluate fractions was assessed by Western blot as described below.

Western blotting—Cell lysates were prepared by lysing cells in ice-cold RIPA buffer (25mM Tris-HCl pH 7.6, 0.15M NaCl, 1% IGEPAL CA-630, 1% sodium deoxycholate, 0.1% SDS) containing 1X protease inhibitors (Thermo Scientific). Lysate was cleared by centrifugation at 20,000 × g for 15 min at 4°C. Samples were boiled in 1X LDS loading buffer (Invitrogen) and 50mM DTT. Proteins were separated by SDS-PAGE using 4–12% Bis-Tris NuPAGE gels (TARBP2, beta tubulin, histone H3) or 3–8% Tris-Acetate NuPAGE gels (METTL3), transferred to 0.2um PVDF (Millipore), blocked using 5% nonfat milk and probed using target-specific antibodies. Bound antibodies were detected using horseradish peroxidase-conjugated secondary antibodies and ECL Western Blotting Substrate (Thermo Scientific), according to the manufacturer's instructions. Antibodies: beta-tubulin (Proteintech 66240–1-Ig), TARBP2 (Proteintech 15753), histone H3 (Proteintech 17168–1-AP), METTL3 (Abcam ab195352).

MeRIP-seq—Biological replicates of control and TARBP2 knockdown cells (MDA-LM2 background) were collected and processed as previously described (Alarcón et al., 2015b). 1×10^7 cells per sample were lysed using LB1 buffer (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% Triton X-100) and 1X protease inhibitor cocktail (Thermo Scientific). The nuclear fraction was then lysed with M-PER buffer (Thermo Scientific) and diluted tenfold in dilution buffer (50 mM Tris-Cl, pH 7.4, 100 mM NaCl) before the immunoprecipitation. Rabbit anti-m6A antibody (Synaptic Systems) and rabbit IgG control bound to protein A dynabeads (Invitrogen) were used for the immunoprecipitations. The immunoprecipitated RNA was eluted with N6-methyladenosine (Sigma-Aldrich), ethanol precipitated and resuspended in water. RNA was barcoded using ScriptSeq V2 kit (Epicentre) and sequenced at Rockefeller University Genomics Core.

Subcellular fractionation—Subcellular fractionation was performed using a low-salt method, based on (Mendez and Stillman, 2000). 1.5×10^7 MDA-LM2 cells were washed in PBS, collected by scraping, and the cell pellet was resuspended in 200µl cold buffer A (10mM HEPES-KOH pH 7.9, 1.5mM MgCl2, 10mM KCl, 340mM sucrose, 10% glycerol, 0.1% Triton X-100, 1mM DTT, 1X halt protease inhibitor cocktail (Thermo Scientific)), and rotated end-over-end 10 minutes at 4°C. Sample was then centrifuged at 1,300 × g 5 minutes at 4°C, and supernatant was removed and clarified by centrifuging at 20,000 × g 10 minutes at 4°C. This supernatant was used as the cytoplasmic fraction. The nuclei were washed with 200µl buffer A without Triton X-100, then centrifuged at 1,300 × g 10 minutes 4°C. Nuclei were then resuspen ded in 100µl buffer B (3mM EDTA, 0.2mM EGTA, 1mM DTT, 1x halt protease inhibitor cocktail (Thermo Scientific)), and rotated end-overend 30 minutes at 4°C. Sample was then centrifuged 1,700 × g 10 minutes at 4°C, and the resulting supernatant was

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

used as the soluble nuclear fraction. The remaining pellet was washed with 200 μl cold phosphate buffered saline, then centrifuged at 1,700 × g 5 minutes at 4°C, and the supernatant discarded. The chromatin p ellet was then DNase treated by adding 44μl H2O, 5μl 10X Turbo DNase buffer, and 2 units of turbo DNase (Invitrogen). The mixture was incubated 10 minutes at 37°C, and was used as the chromatin fraction. Total lysate was prepared by lysing pelleted cells in RIPA (25mM Tris-HCl pH 7.6, 0.15M NaCl, 1% IGEPAL CA-630, 1% sodium deoxycholate, 0.1% SDS) and centrifuging at 20,000 × g 10 minutes to clarify lysate. For Western blotting, 1% of total, and 10% each of the cytoplasmic, soluble nuclear, and chromatin fractions separated by PAGE and detected by immunoblotting as described above.

Reporter construction and reporter splicing assay—To construct an intron retention reporter, the pLX302 backbone (Addgene #25896) was digested with EcoRI and NheI (NEB), and then Gibson assembly was used to add both a synthesized DNA fragment (Genewiz; gAN006, see Table S4 for all oligo sequences used for reporter construction) and the EF1alpha promoter-containing PCR product obtained with primers oAN093 and oAN094 using vector pWPI (Addgene #12254) as the template. The resulting lentiviral vector, pAN30, contains the following elements: 5' LTR, RRE, hPGK promoter, PuroR, cPPT, hEF1alpha promoter, MCS, WPRE, 3' LTR (ΔU3). For cloning the overexpression plasmids, the TARBP2 ORF was PCR amplified with primers oAN377 and oAN378 using pENTR223-TARBP2 (Human ORFeome clone #4636) as template, and mCherry was PCR amplified with primers oAN379 and oAN380 from pLX304-mCherry The resulting PCR products were Gibson assembled with PacI-linearized pAN30.

To generate the intron retention reporters, human HNRNPC intron 1 with flanking exon 1 and 2 sequences (hg38 coordinates chr14:21,263,341–21,269,479) was PCR-amplified with primers oAN381–382, oAN383–385 using human genomic DNA as the template, and cloned downstream of the hEF1alpha promoter in PacI-linearized pAN30 using Gibson assembly. A version of this vector with the intronic TARBP2 binding site deleted (chr14:21,264,116–21,266,455) was also constructed (using primer pairs oAN381–382 and oAN384–385). MDA-LM2 cells stably expressing these reporters were generated by lentiviral delivery of the reporter vectors followed by puromycin selection, as described above. Knockdown of TARBP2 and METTL3 in MDA-LM2 cells was performed by transfecting siRNAs targeting these genes, along with a non-targeting control siRNA, as described above. Total RNA was isolated from these cells 72 hours post-transfection, and the reporter splicing patterns were measured by qRT-PCR using intron-exon and exon-exon spanning primers: pAN30_exon1_fwd, pAN30_intron_rev, and pAN30_exon2_rev (See Table S1).

SRSF1 HITS-CLIP—HITS-CLIP was performed as previously described (Moore et al., 2014). Biological replicates of H1299 cells or biological replicates of MDA-LM2 cells were crosslinked with 400mJ/cm^2 254nm UV. Crosslinked cells were then lysed on ice in low salt buffer (1X PBS, 0.1% SDS, 0.5% sodium deoxycholate, 0.5% IGEPAL CA-630) supplemented with SuperaseIn (Invitrogen) and 1X protease inhibitors. Lysate was then treated with DNase I (Promega) at 37°C for 5 minutes. Lysate was then treated with RNase

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

A (Thermo Scientific), used at a final concentration of 80ng/mL and incubated at 37°C for 5 minutes. Lysate was clarified by spinning at 20,000 × g at 4°C for 20 minutes. The clarified lysate was transferred to protein A dynabeads (Invitrogen) conjugated to anti-SRSF1 antibody (Bethyl) and rotated end-over-end at 4°C for 2 hours. The beads were then washed 2X with low salt buffer, then 2X with high salt buffer (5X PBS, 0.1% SDS, 0.5% sodium deoxycholate, 0.5% IGEPAL CA-630), and then 2X with PNK buffer (50mM Tris pH 7.4, 10mM MgCl2, 0.5% IGEPAL CA-630). The immunoprecipitated protein-RNA complexes were then dephosphorylated on-bead with CIP (NEB), then washed 1X with PNK buffer, 1X with PNK + EGTA buffer (50mM Tris pH 7.4, 20mM EGTA, 0.5% IGEPAL CA-630), then 2X with PNK buffer. T4 RNA ligase (Thermo Scientific) was then used to ligate a 5' ^{32}P-labeled RNA linker (RL3, see Table S5 for sequences of all oligos used for CLIP-seq) to the samples on-bead overnight at 16°C. The samples were then washed 1X with low salt buffer, 1X with high salt buffer, and 2X with PNK buffer. The RNA was then phosphorylated on-bead using PNK (NEB), the beads were subsequently washed 2X with PNK, and eluted by heating in 1X NUPAGE LDS loading buffer at 70°C for 10 minutes at 1,000rpm. The eluates were separated on a 4–12% Bis-Tris NuPAGE gel (Invitrogen) and then transferred to Whatman BA85 nitrocellulose (Sigma). The membrane was then exposed to film to determine the migration of the RNA-protein complexes, and the relevant region was cut from the membrane for library preparation and cut into small pieces.

The RNA from the membrane was then isolated by digesting with 200ul of proteinase K solution (4mg/mL proteinase K (Invitrogen), 100mM Tris pH 7.5, 50mM NaCl, 10mM EDTA) and incubating at 37°C for 20 minutes at 1,000rpm, then adding 200ul PK-urea solution (100mM Tris pH 7.5, 50mM NaCl, 10mM EDTA, 7M urea) and incubating at 37°C for 20 minutes at 1,000rpm, then adding 400ul acid phenol (Sigma) and 130ul chloroform (Sigma) and incubating at 37°C for 20 minutes at 1,000rpm. Tubes were vortexed and spun, and RNA was precipitated from the aqueous layer. The RNA pellet was washed and ligated to the RL5D linker using T4 RNA ligase and incubating overnight at 16°C. The ligation reaction was then treated with DNase I (Promega), extracted with acid phenol chloroform, and the aqueous layer was precipitated.

The RNA was washed and then cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen) and DP3 primer. After the annealing step, the reverse transcription was performed at 50°C for 45 minutes, 55°C for 15 m inutes, then 90°C for 5 minutes. The first round of PCR was carried out using this cDNA, DP3 and DP5 primers and Accuprime Pfx Supermix (Invitrogen) using the following cycle conditions: 1. 95°C 2min, 2. 95°C 20s, 3. 58°C 30s, 4. 68°C 20s. Steps 2–4 were repeated X22–24 5. 68°C 5min. The PCR1 products were gel purified using 10% Urea-TBE PAGE. A second PCR step was then performed to attach Illumina flowcell adapter sequences. PCR2 reactions were carried out using the gel purified PCR1 DNA as template, DSFP5 and DSFP3 primers, Accuprime Pfx Supermix (Invitrogen), and the following cycle conditions: 1. 95°C 2min, 2. 95°C 2 0s, 3. 58°C 30s, 4. 68°C 40s. Steps 2–4 were repeated X6–8. 5. 68°C 5min. These PCR2 products were gel purified using a 2% metaphor agarose (Lonza)/TBE gel. The resulting libraries were sequenced using SSP1 sequencing primer as the custom sequencing primer.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RNA-seq library preparation—Unless otherwise specified below, RNA sequencing libraries were prepared using RNA that had been rRNA depleted using Ribo-Zero Gold (Illumina) followed by ScriptSeq-v2 (Illumina), and sequenced on an Illumina HiSeq2500 or Illumina NextSeq500 instrument at Rockefeller genomic resource center. RNA-seq libraries for expression profiling of MDA-LM2 cells with shRNA-mediated XRN2, EXOSC10, RBM7 or UPF1 knockdown were generated using the QuantSeq 3' mRNA-Seq library prep kit fwd (Lexogen) per the manufacturer's protocol, and sequenced on an Illumina HiSeq4000 at UCSF CAT.

RNA-seq. MeRIP-seq, and CLIP-seq data processing and analysis—To quantify and compare intron retention from the MDA-LM2 TARBP2 knockdown cells, Tophat (v. 2.1.1) was used to map the reads to the human transcriptome (build hg19). Cufflinks and cuffmerge (v.2.2.1) were then used to calculate RPKM. An information-theoretic gene-set enrichment analysis (TEISER; Goodarzi et al., 2012) was used to examine the distribution of TARBP2-bound transcripts among the differentially expressed genes. Transcripts with greater than 0.1 FPKM in the control sample from this experiment were used for all the subsequent analyses. A list of intronic annotations (compatible with MISO) were generated from these expressed transcripts that were bound by TARBP2 based on CLIP-seq data. A randomly selected set of introns with no TARBP2 CLIP tags was used as a representation of background expression. To compare changes in intron retention upon TARBP2 knockdown, we first removed reads mapping to features within introns (miRNAs, snoRNAs, and miscRNAs) using bedtools (intersectBed). MISO (Katz et al., 2010) was then used to estimate changes in intron retention for each of the annotated intronic sequences. The analysis described above was also used to quantify and compare intron retention from the MDA-LM2 WTAP and TPR knockdown RNAseq data, and the MDA-LM2 SRSF1 knockdown RNAseq data.

For comparison of intronic expression in nuclear RNA from control and TARBP2 knockdown cells (MDA-LM2 background), Tophat (v.2.1.1) was used to map the reads to the human transcriptome (build hg19). Cufflinks and cuffmerge (v.2.2.1) were then used to calculate RPKM and cuffdiff was used to compare expression between samples. Reads mapping to features within introns were then removed as described above and the remainder were used to count the number reads mapping to each intronic sequence (featureCounts). DESeq2 was subsequently used to compare expression of TARBP2-bound introns in nuclear RNA from control and TARBP2 knockdown samples. A similar analysis was performed on the background set of introns. The analysis described above was also used to quantify and compare intronic expression in nuclear RNA from control and METTL3 knockdown cells (MDA-LM2 background).

To quantify and compare gene expression and transcript stability from the H1299 cell TARBP2 knockdown RNAseq data, first Tophat (v.2.1.1) was used to map the reads to the human transcriptome (build hg19). Cufflinks, cuffmerge and cuffdiff (v.2.2.1) were then used to calculate RPKM and compare gene expression between samples. REMBRANDTS (Alkallas et al., 2017) was used to compare changes in RNA stability.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To quantify and compare gene expression from the MDA-LM2 cell EXOSC10, XRN2, RBM7 and UPF1 knockdown RNAseq data, STAR (v.2.5.) was used to map reads to the human transcriptome (build hg19). HTSeq-count (v.0.7.) was then used to count reads in each gene, and DESeq2 was used to compare the samples.

To identify m^6A-containing regions from the MeRIP-seq data, first the adapter sequences were removed and quality trimming was performed using Cutadapt. Reads were then mapped to the human genome (build hg19) using Bowtie2 (v.2.3). MDA-LM2 m^6A sites from (Alarcón et al., 2015b) were used as features to count reads in both MeRIP and input samples. EdgeR was then used to compare changes in MeRIP to input in TARBP2 knockdown compared to control cells.

To identify SRSF1-bound sites from the SRSF1 CLIP-seq data, first the adapter sequences were removed and quality trimming was performed Cutadapt. Reads were then mapped to the human genome (build hg19) using BWA (v.0.7.) with the default parameters. The SRSF1 binding sites were then identified using the CTK package (Zhang and Darnell, 2011). For our analysis, the CLIP-derived SRSF1 binding sites from all of the samples were combined to create the list of SRSF1 binding sites.

Quantification and Statistical Analysis

Statistics—Significance reported for enrichment and depletion in heatmaps are based a hypergeometric test. Mutual information (MI) and z-scores are based on an information-theoretic gene-set enrichment analysis (TEISER; Goodarzi et al., 2012) (Figures 1A, 1D, 1F, 2B, 3D, 4A, 4C, 5E, S3G, S3I, S4A, S4B, S4C, S5C, and S5D).

Significance reported for differences in cumulative distribution are based on a Mann-Whitney U test/Wilcoxon signed-rank test (Figures 1B, 2D, 3E, and 3G)

Two-dimensional heatmaps were generated based on a two-dimensional kernel density estimation (bandwidth of 2 and 100 gridpoints). Also reported are the overall correlation between the two gene expression profile. *P* is based on a regression analysis (Figures 2C, 3H, and 4B).

Significance reported for transcript and binding site overlaps as visualized in Venn diagrams are based on a hypergeometric test (Figures 3A, 3C, 3I S1C, S2C, S3C, and S3H).

Differences in gene expression as measured by qRT-PCR were tested for significance with a one-tailed U-test (Figures 5C, 7A, S3B, S3D, S3E, S6C, and S7C). For Figure 3F, an unpaired t-test with Welch's correction was used to calculate significance.

Significance reported for *in vivo* lung colonization and lung xenograft assay is based on both two-way ANOVA (i.e. with time as a continuous co-variate) as a function of time and sub-line identity and also with a one-tailed Mann Whitney U comparison of area under the curve for each mouse (Figures 5F and 6A). For orthotopic lung injections (Figures 5G–H, and 6C), one-tailed U test was used to compare radiance between cohorts. Graphs were generated using Graphpad Prism software, images of bioluminescence signal in mouse experiments are from Living Image software (Perkin Elmer). (Figures 5F, 5G, 5H, 6A and 6C).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Significance reported for gene expression differences between tumor and matched control samples was calculated using a Wilcoxon test (Figures 5A, 6B, and 6D).

Significance reported for Kaplan Meier survival curves was calculated using a log-rank test (Figures 5B and 7C).

Significance reported for scatter plots is based on regression analysis (Figures S3F, S6D, and S7D).

Significance reported for cell proliferation assays is based on fitting an exponential model to fit a growth rate for each sample $(ln(N_{t-1}/N_1) = rt$ where t is measured in days). To compare samples, a model with both time and knockdown state (with or without knockdown) as covariates (with interaction term) was used to assess the significance of variations due to gene knockdown (Figures S5E and S6B).

For all figures: $*P<0.05$, $**P<0.01$, $***P<0.001$.

Data and Software Availability

All sequencing data have been deposited in the GEO database under the accession number GSE130894. No software was generated for this project.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We are grateful to Sohail Tavazoie and Saeed Tavazoie for reading the earlier versions of this manuscript. We acknowledge the UCSF Center for Advanced Technology (CAT) and the Rockefeller Genomics Resource Center for high-throughput sequencing and other genomic analyses and the UCSF Nikon Imaging Center for microscopy. We thank Byron Hann and the Preclinical Therapeutics core as well as the Laboratory Animal Resource Center (LARC) at UCSF. We thank Dr. Trever Bivona for the gift of the H1975-luciferase cell line. We are also grateful for the genomic data contributed by the TCGA Research Network, including donors and researchers. We acknowledge support from our colleagues at the Helen Diller Family Comprehensive Cancer Center. This work was supported by grants from the NIH to H.G (R00CA194077, R01GM123977) and to D.R. (R01CA140456, and R01CA154916). L.F. was supported by NIH training grant T32CA108462-15, Y.X. was supported by a Damon Runyon Postdoctoral Fellowship. and S.Z. was supported through the HHMI Medical Research Fellowship.

REFERENCES

Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, and Tavazoie SF (2015a). HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. Cell 162, 1299–1308. [PubMed: 26321680]

Alarcón CR, Lee H, Goodarzi H, Halberg N, and Tavazoie SF (2015b). N6-methyladenosine marks primary microRNAs for processing. Nature 519, 482–485. [PubMed: 25799998]

Alkallas R, Fish L, Goodarzi H, and Najafabadi HS (2017). Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. Nat. Commun 8, 909. [PubMed: 29030541]

Beer DG, Kardia SLR, Huang C-C, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, et al. (2002). Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat. Med 8, 816–824. [PubMed: 12118244]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Bergeron D, Pal G, Beaulieu YB, Chabot B, and Bachand F (2015). Regulated Intron Retention and Nuclear Pre-mRNA Decay Contribute to PABPN1 Autoregulation. Mol. Cell. Biol 35, 2503–2517. [PubMed: 25963658]

Bryant DM, Datta A, Rodríguez-Fraticelli AE, Peränen J, Martín-Belmonte F, and Mostov KE (2010). A molecular network for de novo generation of the apical surface and lumen. Nat. Cell Biol 12, 1035–1045. [PubMed: 20890297]

Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, and Shiekhattar R (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature 436, 740–744. [PubMed: 15973356]

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, and Mann M (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics MCP 13, 2513–2526. [PubMed: 24942700]

Coyle JH, Bor Y-C, Rekosh D, and Hammarskjold M-L (2011). The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. RNA N. Y. N 17, 1344–1356.

Dai Y, Wang M, Wu H, Xiao M, Liu H, and Zhang D (2016). Loss of FOXN3 in colon cancer activates beta-catenin/TCF signaling and promotes the growth and migration of cancer cells. Oncotarget 8, 9783–9793.

Das S, and Krainer AR (2014). Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. Mol. Cancer Res. MCR 12, 1195–1204. [PubMed: 24807918]

Du H, Zhao Y, He J, Zhang Y, Xi H, Liu M, Ma J, and Wu L (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun 7, 12626. [PubMed: 27558897]

Du P, Kibbe WA, and Lin SM (2008). lumi: a pipeline for processing Illumina microarray. Bioinforma. Oxf. Engl 24, 1547–1548.

Edupuganti RR, Geiger S, Lindeboom RGH, Shi H, Hsu PJ, Lu Z, Wang S-Y, Baltissen MPA, Jansen PWTC, Rossa M, et al. (2017). N(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis. Nat. Struct. Mol. Biol 24, 870–878. [PubMed: 28869609]

Falk S, Finogenova K, Melko M, Benda C, Lykke-Andersen S, Jensen TH, and Conti E (2016). Structure of the RBM7-ZCCHC8 core of the NEXT complex reveals connections to splicing factors. Nat. Commun 7, 13573. [PubMed: 27905398]

Gatignol A, Buckler-White A, Berkhout B, and Jeang KT (1991). Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. Science 251, 1597–1600. [PubMed: 2011739]

Goodarzi H, Najafabadi HS, Oikonomou P, Greco TM, Fish L, Salavati R, Cristea IM, and Tavazoie S (2012). Systematic discovery of structural elements governing stability of mammalian messenger RNAs. Nature 485, 264–268. [PubMed: 22495308]

Goodarzi H, Zhang S, Buss CG, Fish L, Tavazoie S, and Tavazoie SF (2014). Metastasis-suppressor transcript destabilization through TARBP2 binding of mRNA hairpins. Nature 513, 256–260. [PubMed: 25043050]

Győrffy B, Surowiak P, Budczies J, and Lanczky A (2013). Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. PloS One 8, e82241. [PubMed: 24367507]

Haibara H, Yamazaki R, Nishiyama Y, Ono M, Kobayashi T, Hokkyo-Itagaki A, Nishisaka F, Nishiyama H, Kurita A, Matsuzaki T, et al. (2017). YPC-21661 and YPC-22026, novel small molecules, inhibit ZNF143 activity in vitro and in vivo. Cancer Sci. 108, 1042–1048. [PubMed: 28192620]

Karanth S, Zinkhan EK, Hill JT, Yost HJ, and Schlegel A (2016). FOXN3 regulates hepatic glucose utilization. Cell Rep. 15, 2745–2755. [PubMed: 27292639]

Katz Y, Wang ET, Airoldi EM, and Burge CB (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat. Methods 7, 1009–1015. [PubMed: 21057496]

Kawatsu Y, Kitada S, Uramoto H, Zhi L, Takeda T, Kimura T, Horie S, Tanaka F, Sasaguri Y, Izumi H, et al. (2014). The combination of strong expression of ZNF143 and high MIB-1 labelling index independently predicts shorter disease-specific survival in lung adenocarcinoma. Br. J. Cancer 110, 2583–2592. [PubMed: 24736586]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Ke S, Pandya-Jones A, Saito Y, Fak JJ, Vågbø CB, Geula S, Hanna JH, Black DL, Darnell JE, and Darnell RB (2017). m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev. 31, 990–1006. [PubMed: 28637692]

Kilchert C, Wittmann S, Passoni M, Shah S, Granneman S, and Vasiljeva L (2015). Regulation of mRNA Levels by Decay-Promoting Introns that Recruit the Exosome Specificity Factor Mmi1. Cell Rep. 13, 2504–2515. [PubMed: 26670050]

Kilchert C, Wittmann S, and Vasiljeva L (2016). The regulation and functions of the nuclear RNA exosome complex. Nat. Rev. Mol. Cell Biol 17, 227–239. [PubMed: 26726035]

Kim I-J, Quigley D, To MD, Pham P, Lin K, Jo B, Jen K-Y, Raz D, Kim J, Mao J-H, et al. (2013). Rewiring of human lung cell lineage and mitotic networks in lung adenocarcinomas. Nat. Commun 4, 1701. [PubMed: 23591868]

Kim Y, Yeo J, Lee JH, Cho J, Seo D, Kim J-S, and Kim VN (2014). Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of TRBP. Cell Rep. 9, 1061–1074. [PubMed: 25437560]

Knuckles P, Carl SH, Musheev M, Niehrs C, Wenger A, and Bühler M (2017). RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat. Struct. Mol. Biol 24, 561–569. [PubMed: 28581511]

Krull S, Thyberg J, Björkroth B, Rackwitz H-R, and Cordes VC (2004). Nucleoporins as components of the nuclear pore complex core structure and Tpr as the architectural element of the nuclear basket. Mol. Biol. Cell 15, 4261–4277. [PubMed: 15229283]

Laraki G, Clerzius G, Daher A, Melendez-Peña C, Daniels S, and Gatignol A (2008). Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions. RNA Biol. 5, 92–103. [PubMed: 18421256]

Lin S, Choe J, Du P, Triboulet R, and Gregory RI (2016). The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol. Cell 62, 335–345. [PubMed: 27117702]

Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol 10, 93–95. [PubMed: 24316715]

Liu N, Dai Q, Zheng G, He C, Parisien M, and Pan T (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518, 560–564. [PubMed: 25719671]

Macias S, Cordiner RA, Gautier P, Plass M, and Cáceres JF (2015). DGCR8 Acts as an Adaptor for the Exosome Complex to Degrade Double-Stranded Structured RNAs. Mol. Cell 60, 873–885. [PubMed: 26687677]

Méndez J, and Stillman B (2000). Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis. Mol. Cell. Biol 20, 8602–8612. [PubMed: 11046155]

Meola N, Domanski M, Karadoulama E, Chen Y, Gentil C, Pultz D, Vitting-Seerup K, Lykke-Andersen S, Andersen JS, Sandelin A, et al. (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Mol. Cell 64, 520–533. [PubMed: 27871484]

Miki TS, and Großhans H (2013). The multifunctional RNase XRN2. Biochem. Soc. Trans 41, 825–830. [PubMed: 23863139]

Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, Viale A, Olshen AB, Gerald WL, and Massague J (2005). Genes that mediate breast cancer metastasis to lung. Nature 436, 518–524. [PubMed: 16049480]

Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, et al. (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell 124, 1283–1298. [PubMed: 16564017]

Moore MJ, Zhang C, Gantman EC, Mele A, Darnell JC, and Darnell RB (2014). Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. Nat. Protoc 9, 263–293. [PubMed: 24407355]

Nasif S, Contu L, and Muhlemann O (2017). Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression. Semin. Cell Dev. Biol

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Pati D, Keller C, Groudine M, and Plon SE (1997). Reconstitution of a MEC1-independent checkpoint in yeast by expression of a novel human fork head cDNA. Mol. Cell. Biol 17, 3037–3046. [PubMed: 9154802]

Ping X-L, Sun B-F, Wang L, Xiao W, Yang X, Wang W-J, Adhikari S, Shi Y, Lv Y, Chen Y-S, et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 24, 177–189. [PubMed: 24407421]

Rajanala K, and Nandicoori VK (2012). Localization of nucleoporin Tpr to the nuclear pore complex is essential for Tpr mediated regulation of the export of unspliced RNA. PloS One 7, e29921. [PubMed: 22253824]

Rappsilber J, Mann M, and Ishihama Y (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc 2, 1896–1906. [PubMed: 17703201]

Rindler TN, Stockman CA, Filuta AL, Brown KM, Snowball JM, Zhou W, Veldhuizen R, Zink EM, Dautel SE, Clair G, et al. (2017). Alveolar injury and regeneration following deletion of ABCA3. JCI Insight 2.

Scott KL, and Plon SE (2005). CHES1/FOXN3 interacts with Ski-interacting protein and acts as a transcriptional repressor. Gene 359, 119–126. [PubMed: 16102918]

Shi Y, Zuo D, Wang X, Han M, and Wu Y (2016). shRNA-mediated silencing of TARBP2 inhibits NCI-H1299 non-small cell lung cancer cell invasion and migration via the JNK/STAT3/AKT pathway. Mol. Med. Rep 14, 3725–3730. [PubMed: 27599909]

Shulenin S, Nogee LM, Annilo T, Wert SE, Whitsett JA, and Dean M (2004). ABCA3 gene mutations in newborns with fatal surfactant deficiency. N. Engl. J. Med 350, 1296–1303. [PubMed: 15044640]

Sun J, Li H, Huo Q, Cui M, Ge C, Zhao F, Tian H, Chen T, Yao M, and Li J (2016). The transcription factor FOXN3 inhibits cell proliferation by downregulating E2F5 expression in hepatocellular carcinoma cells. Oncotarget 7, 43534–43545. [PubMed: 27259277]

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43, D447–452. [PubMed: 25352553]

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, and Cox J (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 13, 731–740. [PubMed: 27348712]

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, et al. (2014a). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117–120. [PubMed: 24284625]

Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, Weng X, Chen K, Shi H, and He C (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161, 1388–1399. [PubMed: 26046440]

Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, and Zhao JC (2014b). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat. Cell Biol 16, 191–198. [PubMed: 24394384]

Wong JJ-L, Au AYM, Ritchie W, and Rasko JEJ (2016). Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biology. BioEssays News Rev. Mol. Cell. Dev. Biol 38, 41–49.

Xiao W, Adhikari S, Dahal U, Chen Y-S, Hao Y-J, Sun B-F, Sun H-Y, Li A, Ping X-L, Lai W-Y, et al. (2016). Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol. Cell 61, 507–519. [PubMed: 26876937]

Xu C, Wang X, Liu K, Roundtree IA, Tempel W, Li Y, Lu Z, He C, and Min J (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat. Chem. Biol 10, 927–929. [PubMed: 25242552]

Yap K, Lim ZQ, Khandelia P, Friedman B, and Makeyev EV (2012). Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev. 26, 1209–1223. [PubMed: 22661231]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Zarnack K, Konig J, Tajnik M, Martincorena I, Eustermann S, Stevant I, Reyes A, Anders S, Luscombe NM, and Ule J (2013). Direct Competition between hnRNP C and U2AF65 Protects the Transcriptome from the Exonization of Alu Elements. Cell 152, 453–466. [PubMed: 23374342]

Zhang C, and Darnell RB (2011). Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat. Biotechnol 29, 607–614. [PubMed: 21633356]

Zhang C, Chen Y, Sun B, Wang L, Yang Y, Ma D, Lv J, Heng J, Ding Y, Xue Y, et al. (2017a). m6A modulates haematopoietic stem and progenitor cell specification. Nature 549, 273–276. [PubMed: 28869969]

Zhang S, Zhao BS, Zhou A, Lin K, Zheng S, Lu Z, Chen Y, Sulman EP, Xie K, Bogler O, et al. (2017b). m6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. Cancer Cell 31, 591–606.e6. [PubMed: 28344040]

Zhao X, Yang Y, Sun B-F, Shi Y, Yang X, Xiao W, Hao Y-J, Ping X-L, Chen YS, Wang W-J, et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res. 24, 1403–1419. [PubMed: 25412662]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RESEARCH HIGHLIGHTS

- TARBP2 controls the stability of its target transcripts in the nucleus
- TARBP2 recruits the methyltransferase complex to deposit m^6A marks on transcripts
- m6A marks increase TARBP2 target transcript intron retention and nuclear decay
- TARBP2 promotes lung cancer growth *in vivo* and in clinical datasets

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

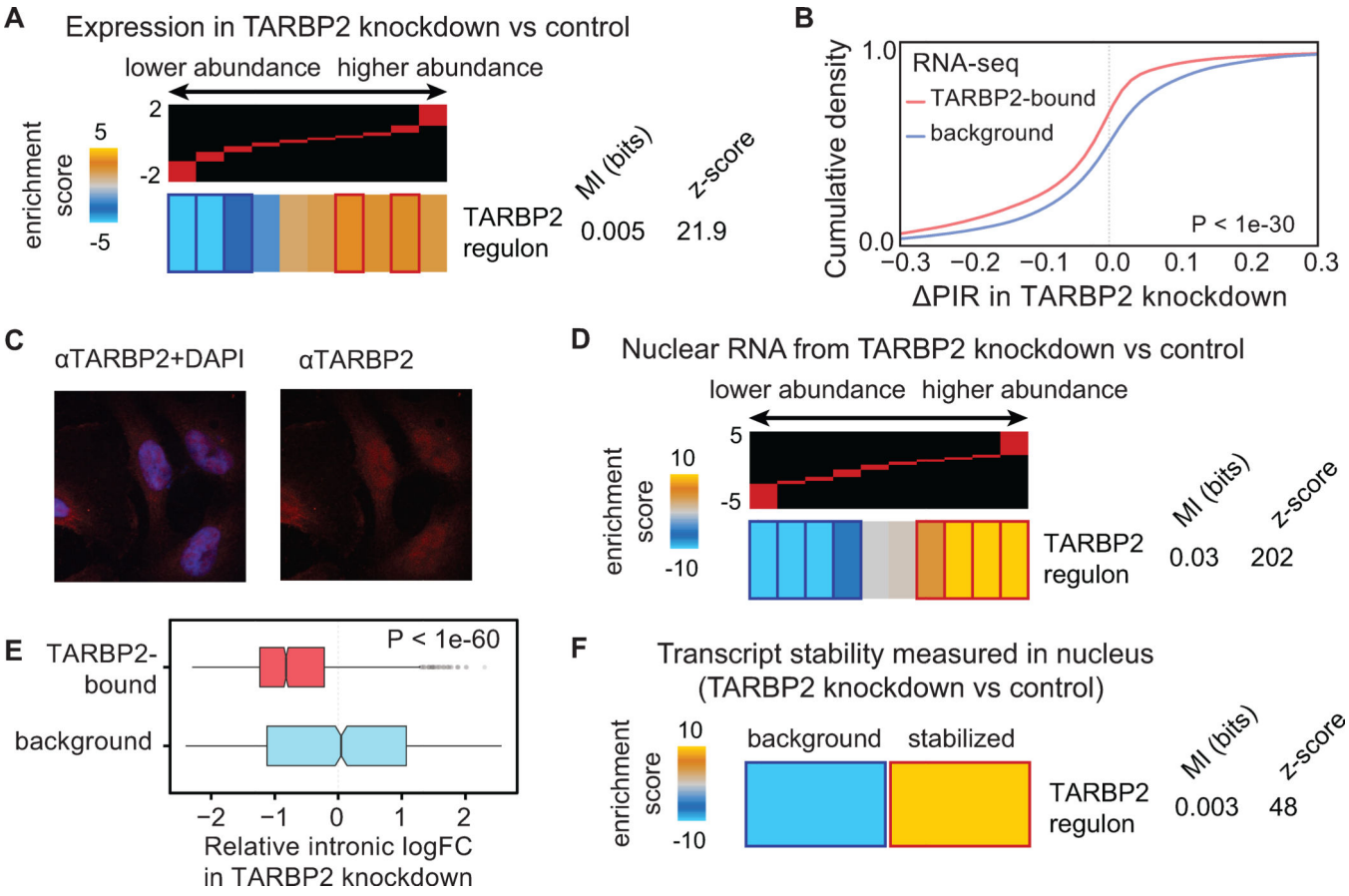


Figure 1 |. Nuclear TARBP2 binding increases pre-mRNA intron retention and decreases transcript stability.

(A) A heatmap showing the enrichment of the TARBP2 regulon among transcripts with higher expression in TARBP2 knockdown compared to control MDA-LM2 breast cancer cells. Genes are sorted based on their expression changes in TARBP2 knockdown cells, from down-regulated (left) to up-regulated (right), and grouped into equally populated bins that are visualized as columns. The red bar on top of every column shows the range of log-fold change values for the genes in its corresponding bin. In the heatmap, high enrichment scores are represented by gold, and correspond to bins with enrichment of TARBP2-bound transcripts, while blue represents depletion of TARBP2-bound transcripts. Statistically significant enrichments and depletions, based on hypergeometric tests, are marked with red and dark-blue borders, respectively. Also included are mutual information (MI) values and their associated z-scores (see STAR Methods). **(B)** Cumulative distribution of changes in percent intron retention (PIR) for TARBP2-bound introns in TARBP2 knockdown relative to control MDA-LM2 breast cancer cells (red line). A background set containing a similar number of introns with no evidence of TARBP2 binding (based on HITS-CLIP) is included as a control (blue line). p calculated using a Mann-Whitney U test. **(C)** Immunofluorescence staining for TARBP2 (red) and staining with DAPI (blue). Shown are z-slices obtained from confocal microscopy imaging of MDA-LM2 cells. Scale bars, 25μm. **(D)** Enrichment of the TARBP2 regulon among the transcripts upregulated in nuclear RNA from TARBP2 knockdown compared to control MDA-LM2 breast cancer cells. **(E)** Relative fold-change of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

percent intron retention for TARBP2-bound introns in nuclear RNAs isolated from TARBP2 knockdown and control MDA-LM2 breast cancer cells. Log fold-changes in intron levels due to TARBP2 knockdown were compared between TARBP2-bound and background introns. The P value was calculated using Mann-Whitney U test. **(F)** Relative stability of nuclear RNA in TARBP2 knockdown compared to control MDA-LM2 cells was measured by microarray. Heatmap displays enrichment of the TARBP2 regulon among transcripts that are stabilized in the nucleus upon TARBP2 knockdown.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

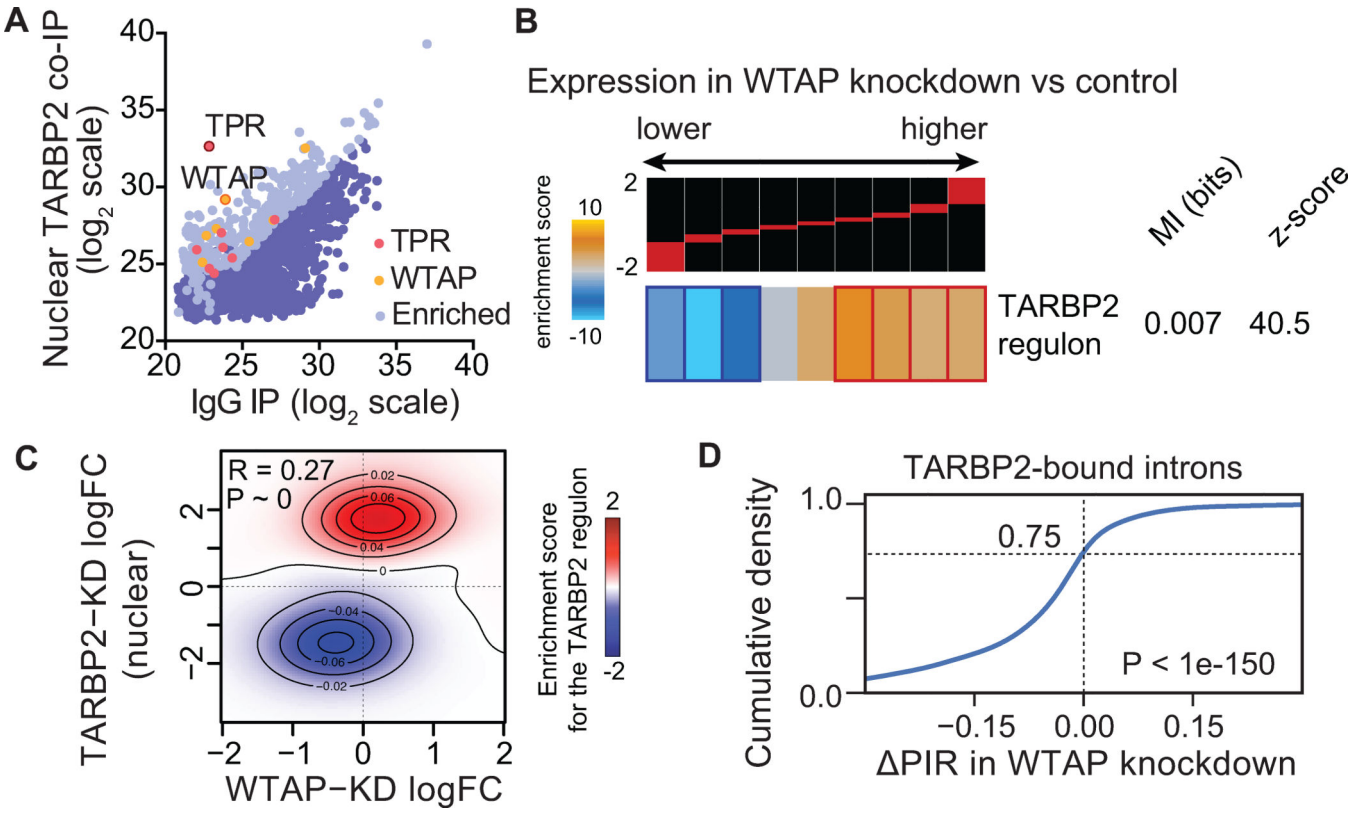


Figure 2 |. TARBP2 regulates intron retention through interactions with N(6)-methyladenosine methyltransferase associated factors.

(A) Scatter plot of mass spectrometry data showing proteins that co-immunoprecipitated with TARBP2 versus control IgG in nuclear lysate. Shown are the average of three replicates across all detected proteins. Proteins enriched in the TARBP2 co-IP samples are shown in light blue. In red are TPR and associated proteins, while WTAP and associated proteins are shown in gold. **(B)** Enrichment of the TARBP2 regulon among transcripts that are upregulated in WTAP knockdown compared to control cells (MDA-LM2 background). Also included are the mutual information value (MI) and the associated *z*-score. **(C)** Two-dimensional heatmap showing enrichment of TARBP2-bound transcripts in the group of transcripts that is upregulated in both TARBP2 and WTAP knockdown cells. Red indicates TARBP2 regulon enrichment, blue depletion. **(D)** Cumulative distribution of changes in percent intron retention (PIR) in WTAP knockdown relative to control MDA-LM2 breast cancer cells. Transcripts not bound by TARBP2 were not included as they may be regulated by WTAP independently of TARBP2. p based on a Wilcoxon signed-rank test.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

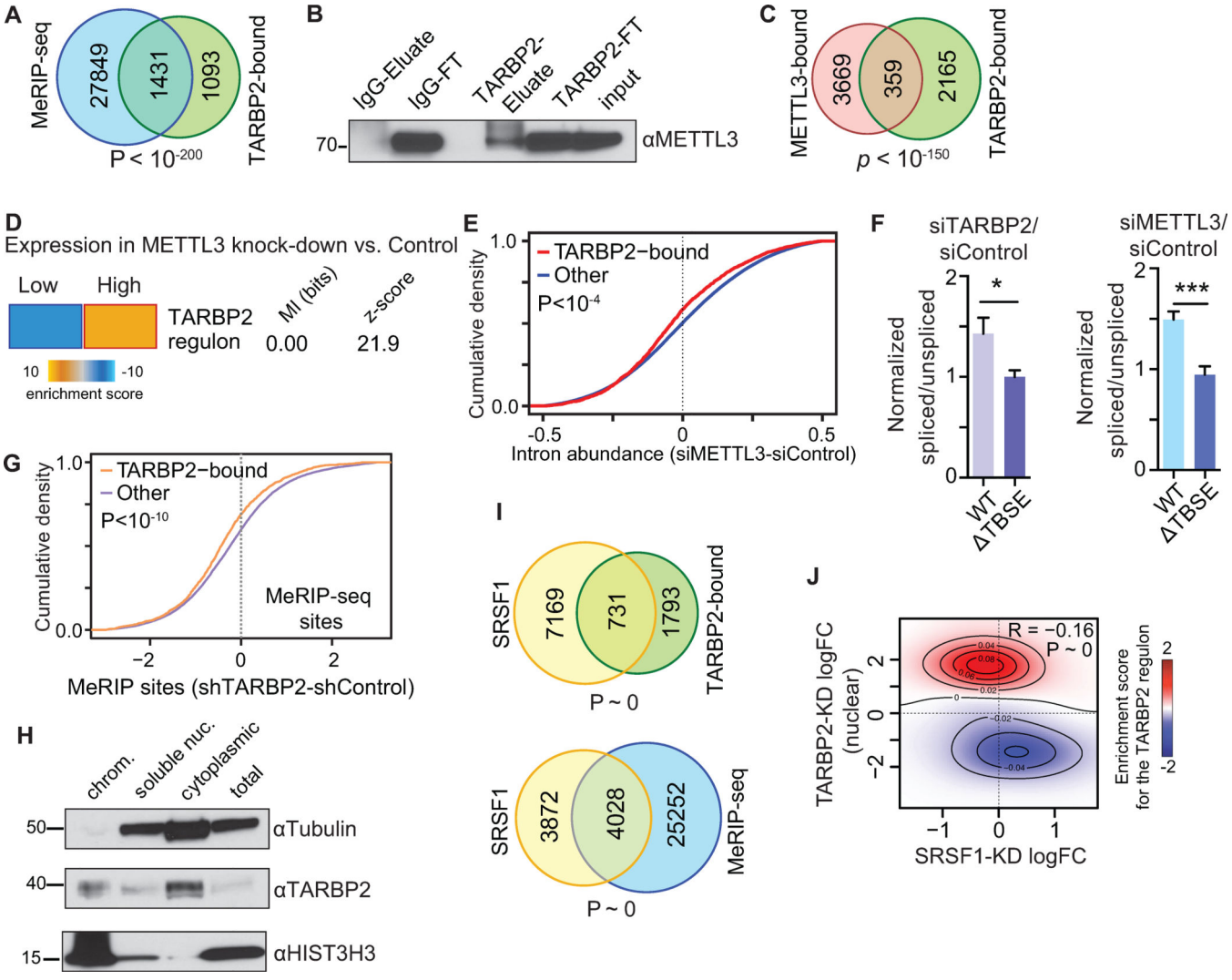


Figure 3 |. TARBP2-dependent m^{6}A modification increases intron retention.

(A) Venn diagram showing significant overlap between introns and flanking exons bound by TARBP2 and those containing m^{6}A marks in MDA-LM2 breast cancer cells. **(B)** Western blot for METTL3 for input, flowthrough and eluate from TARBP2 and IgG immunoprecipitations in MDA-LM2 lysate. **(C)** Venn diagram showing significant overlap between introns and flanking exons bound by TARBP2 and those bound by METTL3 based on a previously published METTL3 PAR-CLIP data (Liu et al., 2014). **(D)** Heatmap showing enrichment of the TARBP2 regulon among transcripts upregulated upon METTL3 knockdown relative to control MDA-LM2 breast cancer cells (data from Alarcón et al., 2015b). The associated mutual information (MI) and *z*-score are shown. (**E**) RNA-seq was performed on nuclear RNA from MDA-LM2 METTL3 knockdown and control cells. Cumulative distribution graph showing a METTL3-dependent decrease in intron expression for TARBP2-bound introns compared to those without evidence of TARBP2 binding. (**F**) qRT-PCR was used to measure the relative levels of spliced reporter RNA in TARBP2 and METTL3 knockdown compared to control MDA-LM2 cells expressing either a reporter intron with the TARBP2 binding site deleted or a reporter with the wild-type intron.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Experiment was performed in biological triplicates with two independent siRNAs targeting TARBP2 and METTL3 (n = 2×3), and performed once in biological triplicate for siCTRL-transfected cells (n = 3). **(G)** MeRIP-seq was performed in MDA-LM2 TARBP2 knockdown and control cells. Cumulative distribution graph showing introns bound by TARBP2 have significantly reduced methylation upon TARBP2 knockdown compared to introns not bound by TARBP2. **(H)** Cytoplasmic, soluble nuclear and chromatin-associated protein fractions were collected from MDA-LM2 cells, and western blotting was used to detect TARBP2, tubulin (cytoplasmic) and histone H3 (nuclear). **(I)** Venn diagram showing significant overlap between introns and flanking exons bound by SRSF1 and TARBP2 (CLIP-seq) and overlap between intron and flanking exons bound by SRSF1 in MDA-LM2 and H1299 cells (CLIP-seq) and those that contain an m^6A mark (MeRIP-seq). **(H)** Two-dimensional heatmap showing enrichment of TARBP2-bound transcripts in the group of transcripts that is upregulated in both TARBP2 knockdown cells and downregulated in SRSF1 knockdown cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

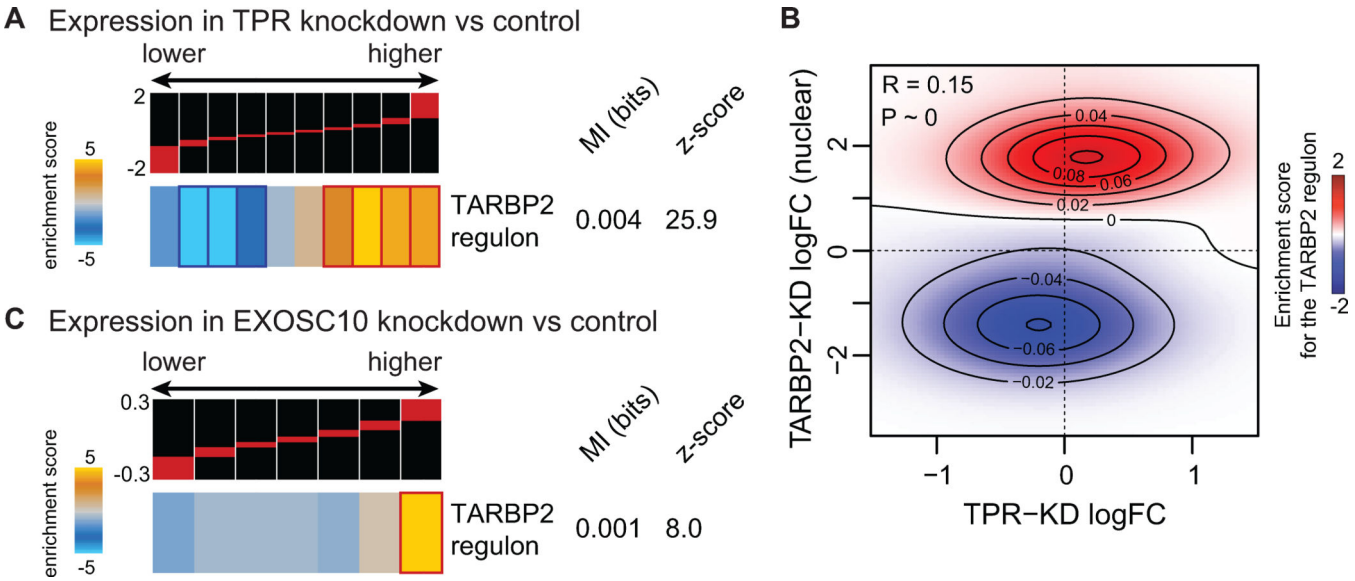


Figure 4 |. RNA processing and nuclear exosome factors control degradation of the TARBP2 regulon.

(A) Heatmap showing enrichment of the TARBP2 regulon among genes upregulated in TPR knockdown compared to control MDA-LM2 breast cancer cells. **(B)** Two-dimensional heatmap showing enrichment of TARBP2-bound transcripts in the group of transcripts that is upregulated in both TARBP2 knockdown cells and TPR knockdown cells. Red indicates TARBP2 regulon enrichment, blue depletion. (**C**) Heatmap showing enrichment of the TARBP2 regulon among transcripts upregulated in EXOSC10 knockdown relative to control MDA-LM2 breast cancer cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

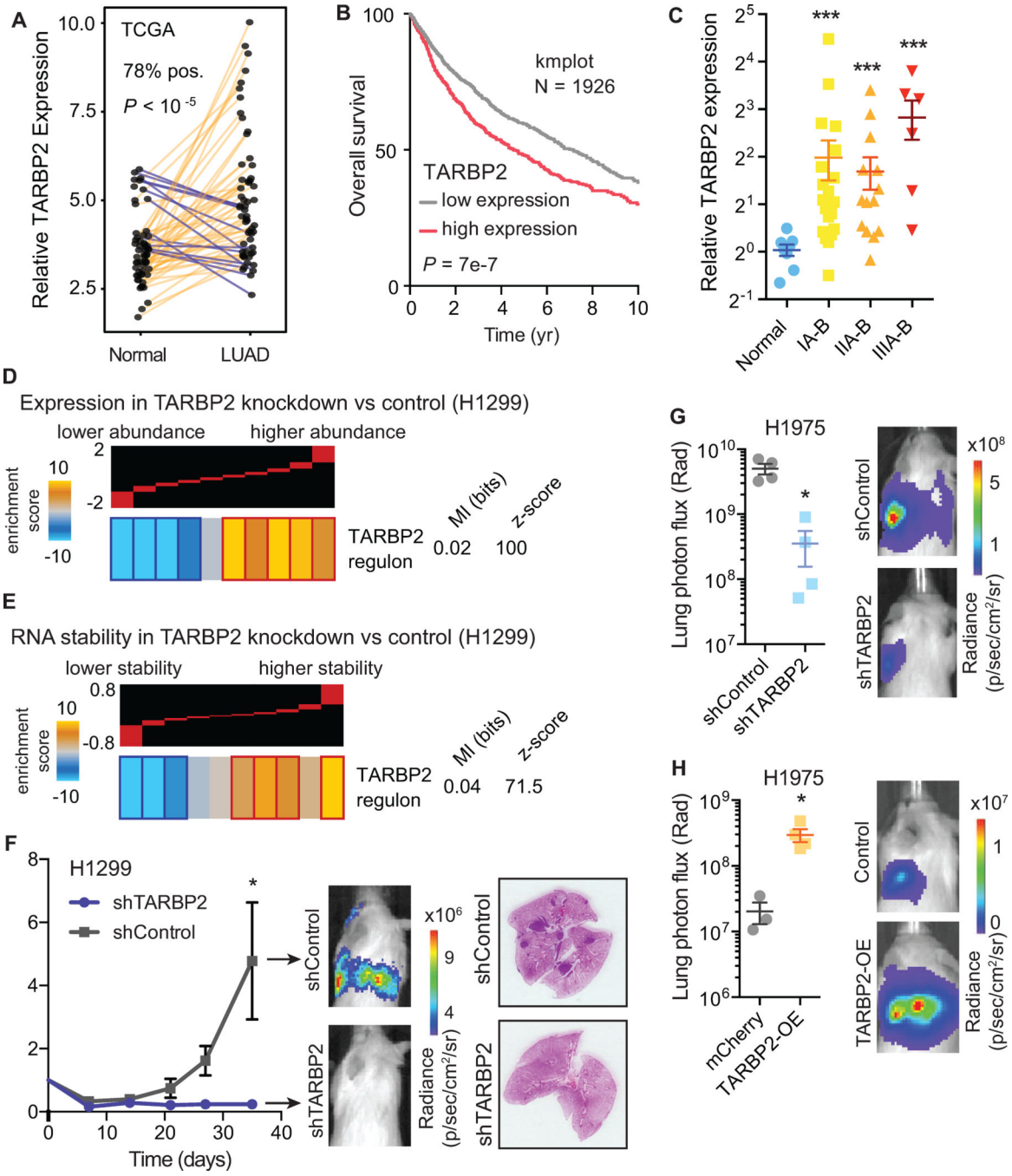


Figure 5 |. TARBP2 promotes lung cancer *in vivo* and is associated with clinical outcome in lung cancer.

(A) Upregulation of TARBP2 mRNA in lung adenocarcinoma (LUAD) samples relative to their matched control in The Cancer Genome Atlas (TCGA) dataset. p was calculated using Wilcoxon tests. **(B)** Kaplan-Meier survival curve showing overall survival of non-small cell lung cancer patients as a function of TARBP2 expression (Győrffy et al., 2013). p calculated using a log-rank test. **(C)** Relative mRNA levels of TARBP2 in normal lung tissue and lung adenocarcinoma stages I-III (N=96) measured using qRT-PCR. p calculated using Mann

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Whitney U test. **(D)** Heatmap showing enrichment of the TARBP2 regulon among transcripts upregulated in H1299 lung cancer cells with TARBP2 knockdown compared to control cells. The associated mutual information (MI) and z-score are also shown. **(E)** TARBP2-dependent transcript stability was inferred in TARBP2 knockdown and control H1299 lung cancer cells by using REMBRANTS (Alkallas et al., 2017) to analyze RNA-seq data from these cell lines. Heat map shows TARBP2-bound transcripts are enriched among transcripts with increased stability in TARBP2 knockdown H1299 lung cancer cells. **(F)** Plots show lung bioluminescence signal over time in mice injected via tail vein with H1299 lung cancer cells expressing a TARBP2 targeting shRNA or a control shRNA. Representative H&E lung histology is shown and the bioluminescence signal from the endpoint is plotted; n = 5 mice per cohort. **(G)** Orthotopic xenografts of H1975 lung cancer cells expressing a TARBP2 targeting shRNA or a control shRNA were performed. Plots show lung bioluminescence signal in mice at day 17 post-injection; n = 4 mice per cohort. **(H)** Orthotopic xenografts of H1975 lung cancer cells stably overexpressing TARBP2 or mCherry were performed. Plots show lung bioluminescence signal in mice at day 10 post-injection; n = 3–4 mice per cohort.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

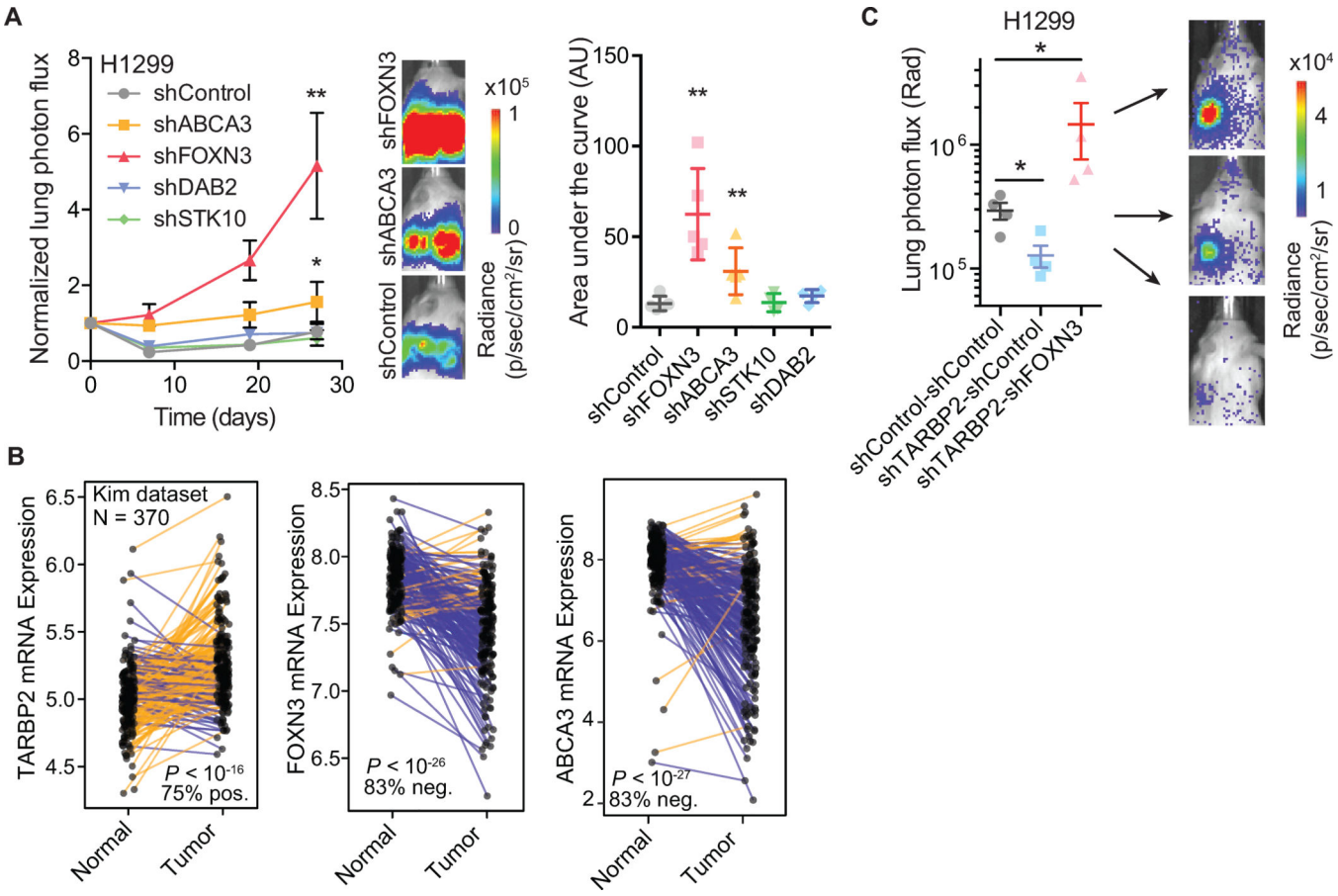


Figure 6 |. FOXN3 acts downstream of TARBP2 to promote lung cancer.

(A) Plots show lung bioluminescence signal over time in mice injected via tail vein with H1299 lung cancer cells expressing shRNAs targeting ABCA3, FOXN3, DAB2, STK10, or a control shRNA; n = 4–5 mice per cohort. (**B**) Relative TARBP2, ABCA3 and FOXN3 mRNA expression in paired normal and lung cancer samples (Kim et al., 2013); p-value was calculated using the Wilcoxon test. (**C**) Orthotopic xenografts of H1299 lung cancer cells expressing control shRNAs, shTARBP2 and a control shRNA, or shTARBP2 and shFOXN3 were performed. Plots show lung bioluminescence signal at day 17 post-injection; n = 4 mice per cohort.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

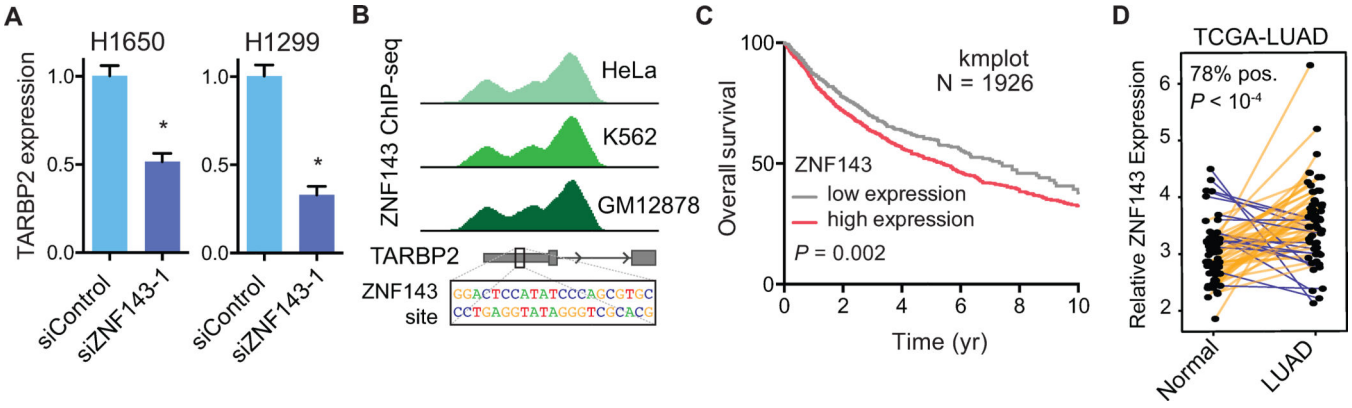


Figure 7 |. ZNF143 regulates TARBP2 expression in breast and lung cancer.

(A) Relative TARBP2 mRNA levels were measured by qRT-PCR in H1650 and H1299 lung cancer cells transfected with siRNAs targeting ZNF143 or a control siRNA; n = 3. **(B)** ENCODE ChIP-seq tracks from three different cell lines that show evidence of ZNF143 binding to the promoter region of TARBP2. These ChIP peaks are also located at a strong sequence match to the ZNF143 consensus motif. **(C)** Kaplan-Meier survival curve showing overall survival of non-small cell lung cancer patients as a function of ZNF143 expression (Győrffy et al., 2013). p was calculated using a log-rank test. **(D)** Relative ZNF143 mRNA expression in paired normal and tumor samples from TCGA-LUAD dataset (p based on a Wilcoxon test).