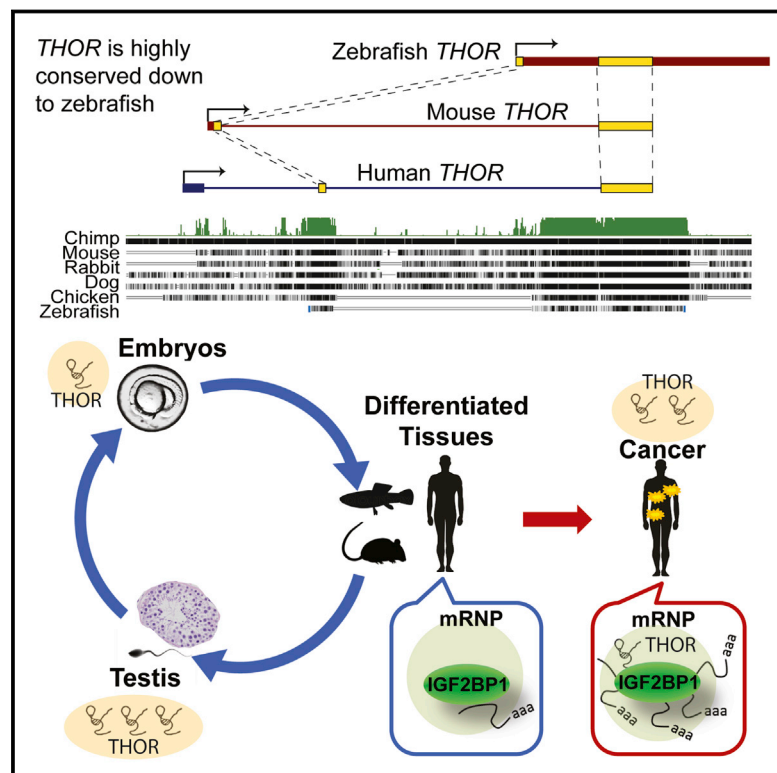# Oncogenic Role of *THOR*, a Conserved Cancer/Testis Long Non-coding RNA

Graphical Abstract

Authors

Yasuyuki Hosono, Yashar S. Niknafs, John R. Prensner, ..., Felix Y. Feng, Weibin Zhou, Arul M. Chinnaiyan

Correspondence

arul@med.umich.edu

In Brief

An ultraconserved lncRNA promotes oncogenesis.

Highlights

- *THOR* is a conserved lncRNA expressed in cancers and in normal testis

- *THOR* functions via a conserved interaction with IGF2BP1, stabilizing its targets

- Oncogenicity of *THOR* is corroborated in a zebrafish model

CellPress

Article

Oncogenic Role of *THOR*, a Conserved Cancer/Testis Long Non-coding RNA

Yasuyuki Hosono,[1,14] Yashar S. Niknafs,[1,2,14] John R. Prensner,[1,3] Matthew K. Iyer,[1,4] Saravana M. Dhanasekaran,[1,5] Rohit Mehra,[1,5] Sethuramasundaram Pitchiaya,[1] Jean Tien,[1] June Escara-Wilke,[13] Anton Poliakov,[1] Shih-Chun Chu,[1] Sahal Saleh,[1] Keerthana Sankar,[1] Fengyun Su,[1] Shuling Guo,[6] Yuanyuan Qiao,[1] Susan M. Freier,[6] Huynh-Hoa Bui,[6] Xuhong Cao,[1] Rohit Malik,[1,5] Timothy M. Johnson,[7,8] David G. Beer,[8,9] Felix Y. Feng,[1,8,10] Weibin Zhou,[11,15] and Arul M. Chinnaiyan[1,2,5,8,12,13,16,*]
[1]Michigan Center for Translational Pathology
[2]Department of Cellular and Molecular Biology
University of Michigan, Ann Arbor, MI, USA
[3]Department of Pediatrics, Boston Children's Hospital, Boston, MA, USA
[4]Department of Computational Medicine and Bioinformatics
[5]Department of Pathology
University of Michigan, Ann Arbor, MI, USA
[6]Ionis Pharmaceuticals, Carlsbad, CA, USA
[7]Department of Dermatology, University of Michigan Medical School, Ann Arbor, MI, USA
[8]Comprehensive Cancer Center
[9]Section of Thoracic Surgery, Department of Surgery
[10]Department of Radiation Oncology
[11]Department of Pediatrics and Communicable Diseases
[12]Howard Hughes Medical Institute
[13]Department of Urology
University of Michigan, Ann Arbor, MI, USA
[14]These authors contributed equally
[15]Present address: Center for Human Disease Modeling, Duke University, 300 N. Duke Street, Durham, NC 27701, USA
[16]Lead Contact
*Correspondence: arul@med.umich.edu
https://doi.org/10.1016/j.cell.2017.11.040

SUMMARY

Large-scale transcriptome sequencing efforts have vastly expanded the catalog of long non-coding RNAs (lncRNAs) with varying evolutionary conservation, lineage expression, and cancer specificity. Here, we functionally characterize a novel ultra-conserved lncRNA, *THOR* (ENSG00000226856), which exhibits expression exclusively in testis and a broad range of human cancers. *THOR* knockdown and overexpression in multiple cell lines and animal models alters cell or tumor growth supporting an oncogenic role. We discovered a conserved interaction of *THOR* with IGF2BP1 and show that *THOR* contributes to the mRNA stabilization activities of IGF2BP1. Notably, transgenic *THOR* knockout produced fertilization defects in zebrafish and also conferred a resistance to melanoma onset. Likewise, ectopic expression of human *THOR* in zebrafish accelerated the onset of melanoma. *THOR* represents a novel class of functionally important cancer/testis lncRNAs whose structure and function have undergone positive evolutionary selection.

INTRODUCTION

Long non-coding RNAs (lncRNAs) have emerged as an abundant and functionally diverse species of ncRNA (Iyer et al., 2015; Ulitsky and Bartel, 2013). Despite their striking prevalence in the transcriptome and countless efforts to interrogate their function, our understanding of the function of the vast majority of lncRNAs remains anecdotal, making their classification particularly challenging (St. Laurent et al., 2015). Novel classes of lncRNAs continue to be identified with categorization criteria largely related to their functional roles and conservation (Iyer et al., 2015; Kim et al., 2010; Prensner and Chinnaiyan, 2011; Ulitsky and Bartel, 2013).

Although the general level of conservation of lncRNAs has been contentious (Cabili et al., 2011; Iyer et al., 2015), there is a clear subclass of lncRNAs that are highly conserved, many of which possess "ultraconserved" regions (i.e., at least 200 bp of nearly perfect vertebrate conservation) (Calin et al., 2007; Hudson et al., 2013; Ulitsky et al., 2011). While conservation is a trait highly suggestive of functional relevance in cells, it also permits the characterization and mechanistic investigation of lncRNAs in model organisms (Sauvageau et al., 2013; Ulitsky et al., 2011).

Here, we define a novel class of lncRNA with normal tissue expression limited to the testis and widespread expression in multiple cancer types. This cancer/testis expression pattern is
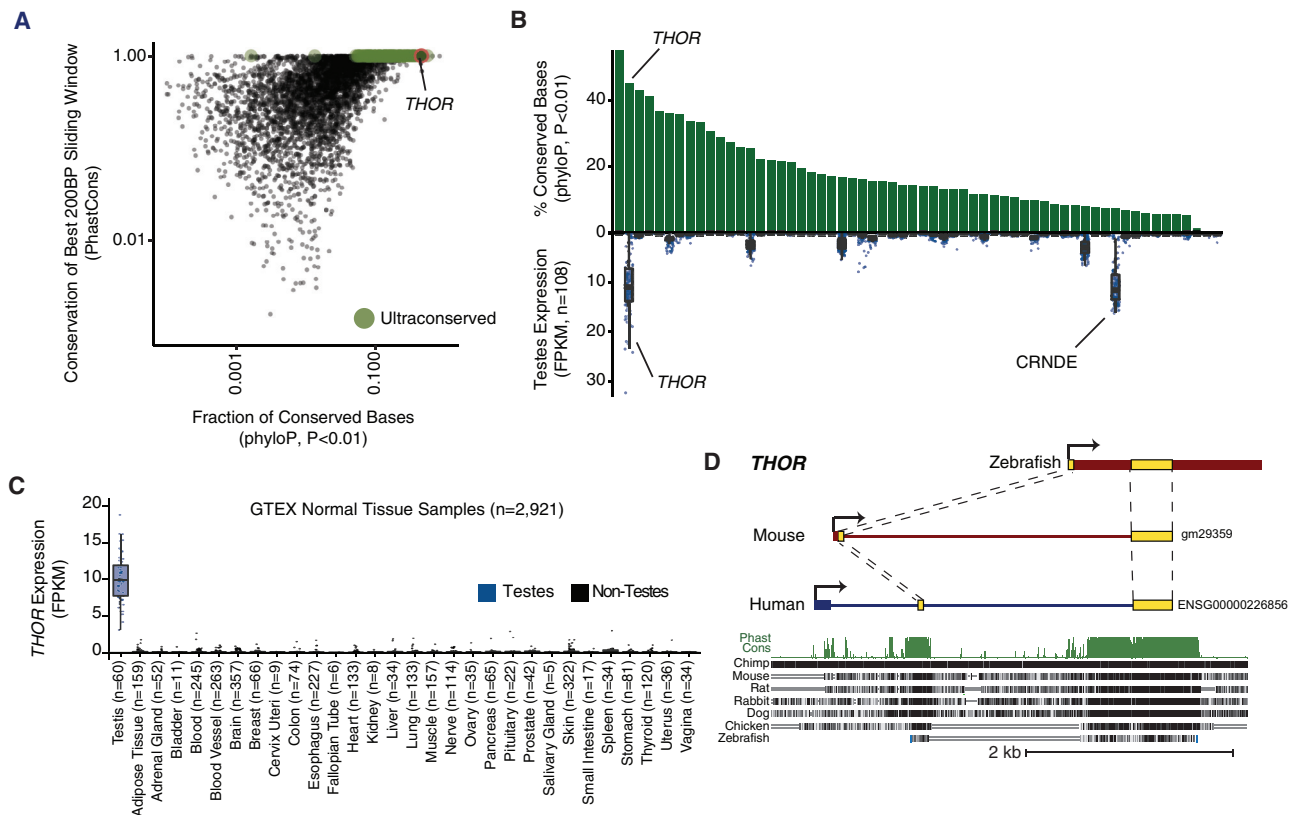
Figure 1. *THOR* Is a Conserved Testis lncRNA

(A) Scatterplot of base-wise transcript conservation levels (x axis) and average conservation for the best 200-bp window (y axis) for all intergenic transcripts expressed at 1 FPKM or more in the top 1% of the Cancer Genome Atlas (TCGA) samples. Green points indicate transcripts with 200-bp windows that meet the criteria for ultraconserved regions (STAR Methods).

(B) Dual plot depicting the fraction of conserved bases (top, green; [A], x axis) and the expression across testes RNA-seq samples (bottom, blue) for all ultraconserved lncRNAs identified in (A).

(C) Expression in FPKM of *THOR* among the GTEx normal tissue RNA-seq dataset.

(D) Genome browser depiction of *THOR* and its conserved analogs in mouse and zebrafish. *THOR* is annotated in the mouse as gm29359. Multiz alignment of multiple vertebrate species depicted as well as the per base PhastCons conservation score.

See also Figures S1 and S2 and Tables S1 and S2.

characteristic of cancer/testis antigens, a well-defined class of proteins that has been suggested as targets for cancer therapy (Fratta et al., 2011; Scanlan et al., 2002; Simpson et al., 2005). Here we characterize *THOR* (testis-associated highly conserved oncogenic long non-coding RNA) and investigate its role in oncogenesis and testis physiology, identifying an evolutionarily conserved functional interaction with IGF2 mRNA-binding proteins (IGF2BPs). Moreover, we leverage the sequence conservation of *THOR* to generate transgenic zebrafish models that shed light on its function in the testis and in cancers.

RESULTS

Discovery of *THOR*, a Conserved lncRNA Expressed in the Testis

In a recent large-scale RNA sequencing (RNA-seq) analysis, we comprehensively profiled the human transcriptome, discovering tens of thousands of novel lncRNAs (Iyer et al., 2015). While lncRNAs tend to be less conserved than protein-coding genes

(Figure S1A), and most do not exhibit marked sequence conservation, a subset of conserved lncRNAs does exist (Figure 1A). We measured both the average base-wise conservation of the entire transcript (Figure 1A, x axis) and the level of conservation of the best 200-bp window (Figure 1A, y axis), a metric previously utilized to determine ultraconserved elements (Hudson et al., 2013; Iyer et al., 2015). We identified 82 intergenic ultraconserved lncRNAs with expression of at least 1 FPKM in the top 1% of samples in our tissue RNA-seq compendium (Figures 1A and 1B and Table S1). Despite possessing a 200-bp ultraconserved segment, these lncRNAs possessed varying degrees of base-wise conservation (Figure 1B, top; range, 0.1%–55.4% conserved bases), with *THOR* exhibiting the second highest degree of base-wise conservation. Interestingly, when interrogating the Genotype-Tissue Expression (GTEx) benign tissue RNA-seq dataset (Consortium, 2013; Melé et al., 2015), two of these ultraconserved lncRNAs, *THOR* and *CRNDE*, displayed substantial expression in the testes (Figure 1B, bottom). We focused further analysis on *THOR* due to its testis-specific
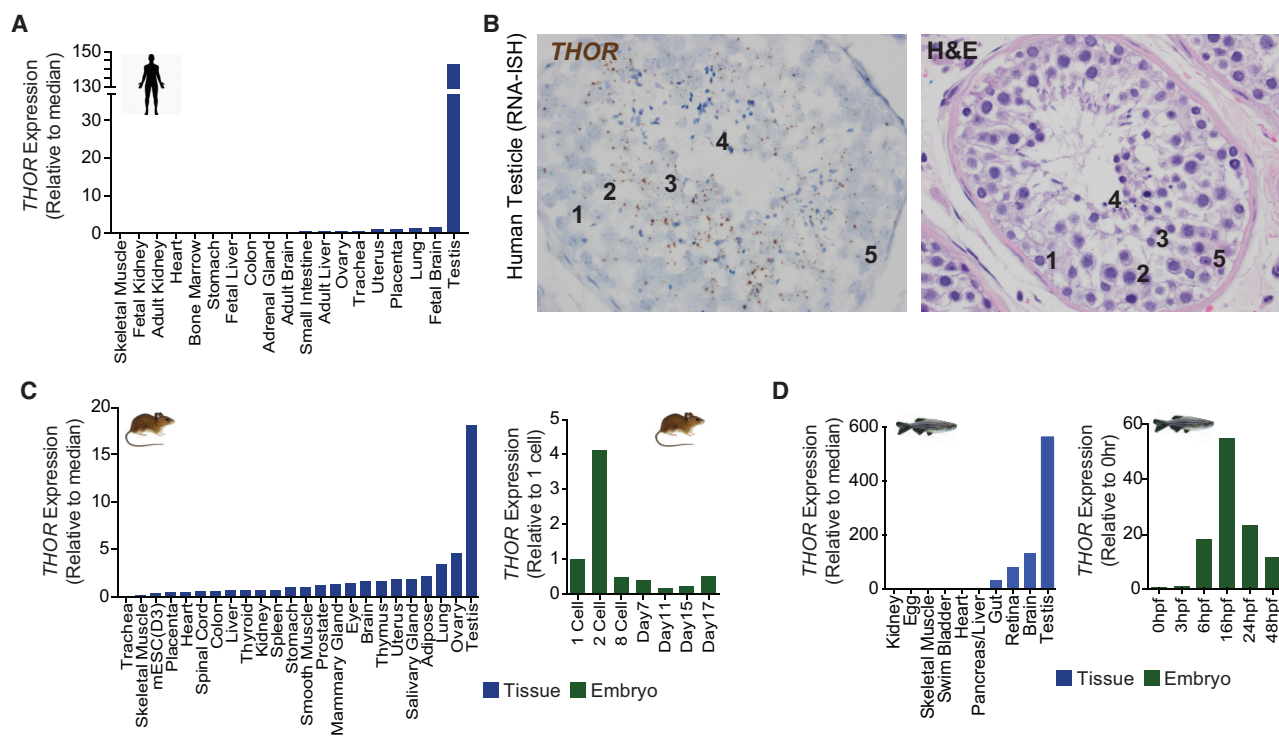
Figure 2. *THOR* Exhibits Testis-Specific Expression

(A) Estimation of *THOR* mRNA expression by qRT-PCR in human adult normal tissue panel.

(B) H&E stain of human testis at high magnification (400×; right), and RNA-ISH of *THOR* in human testis (left). Various cells of the testis are labeled as follows: (1) spermatogonia, (2) spermatocytes, (3) spermatids, (4) mature spermatozoa, and (5) scattered Sertoli cells with a single central prominent nucleolus. *THOR* expression is observed in the spermatid and spermatocyte.

(C) Measurement of mouse *THOR* expression by qRT-PCR on an adult murine tissue panel (left) and embryos (right).

(D) Quantification of zebrafish *THOR* expression by qRT-PCR on a piscine tissue panel (left) and embryos (right).

expression pattern (Figure 1C) compared to the promiscuous expression of *CRNDE* (Figure S1B).

Transcriptional *THOR* Homologs Exist in the Mouse and Zebrafish

Using 5′ and 3′ rapid amplification of cDNA ends (RACE), we identified two isoforms of *THOR*, comprised of either 2 or 3 exons on chromosome 2 (Figure S1C and Table S2). Additionally, existence of the 3-exon isoform of *THOR* was confirmed via northern blotting in the H1299 human lung adenocarcinoma cell line (Figure S2A). While the GENCODE-annotated gene has an additional larger isoform with a downstream exon, expression of this isoform was not detected in any of the cell lines used in this study (Figure S2B), and addition of THOR-targeting siRNA did not alter the expression of the long isoform (Figure S2C). Given its substantial sequence conservation, we set out to identify *THOR* homologs in other species. Utilizing the BLAST-like alignment tool (BLAT) (Kent, 2002) we identified predicted regions in the mouse and zebrafish genome homologous to the human *THOR* (h-*THOR*, Ensemble ID: ENSG00000226856) (Figure 1D). Additionally, we confirmed a 2-exon homolog in mouse *THOR* (m-*THOR*) (Figure S1D and Table S2), and 2 monoexonic isoforms in zebrafish *THOR* (z-*THOR*) (Figure S1E and Table S2). The shorter zebrafish isoform was also detected via northern

blotting (Figure S2D). The conservation of h-*THOR* extends to exons 2 and 3, which are both represented in m-*THOR* and z-*THOR* (Figure 1D). Further characterization of *THOR* confirmed it as a non-coding transcript (Figures S2E–S2G).

THOR Exhibits an Evolutionarily Conserved Expression Pattern in Normal Tissues

To obtain an independent validation of testis-specific *THOR* expression observed in the GTEx RNA-seq data, we performed qRT-PCR with cDNA derived from various normal human tissues, observing a similarly testis-specific expression pattern (Figure 2A). Moreover, RNA *in situ* hybridization (ISH) of human testis tissue using h-*THOR* specific probes revealed an enrichment of *THOR* testis expression in the spermatocyte and spermatid (Figure 2B), but not in surrounding tissue (Figures S2H and S2I). This expression pattern for *THOR* is interestingly a similar expression pattern reported for cancer/testis antigens not found on the X chromosome (Simpson et al., 2005; Tapparel et al., 2003). Querying additional RNA from tissue panels in the mouse and zebrafish identified testis-specific expression for both m-*THOR* (Figure 2C) and z-*THOR* (Figure 2D). Additionally, we also observed elevated *THOR* expression during the early development of both the mouse and zebrafish (Figures 2C and 2D, right).
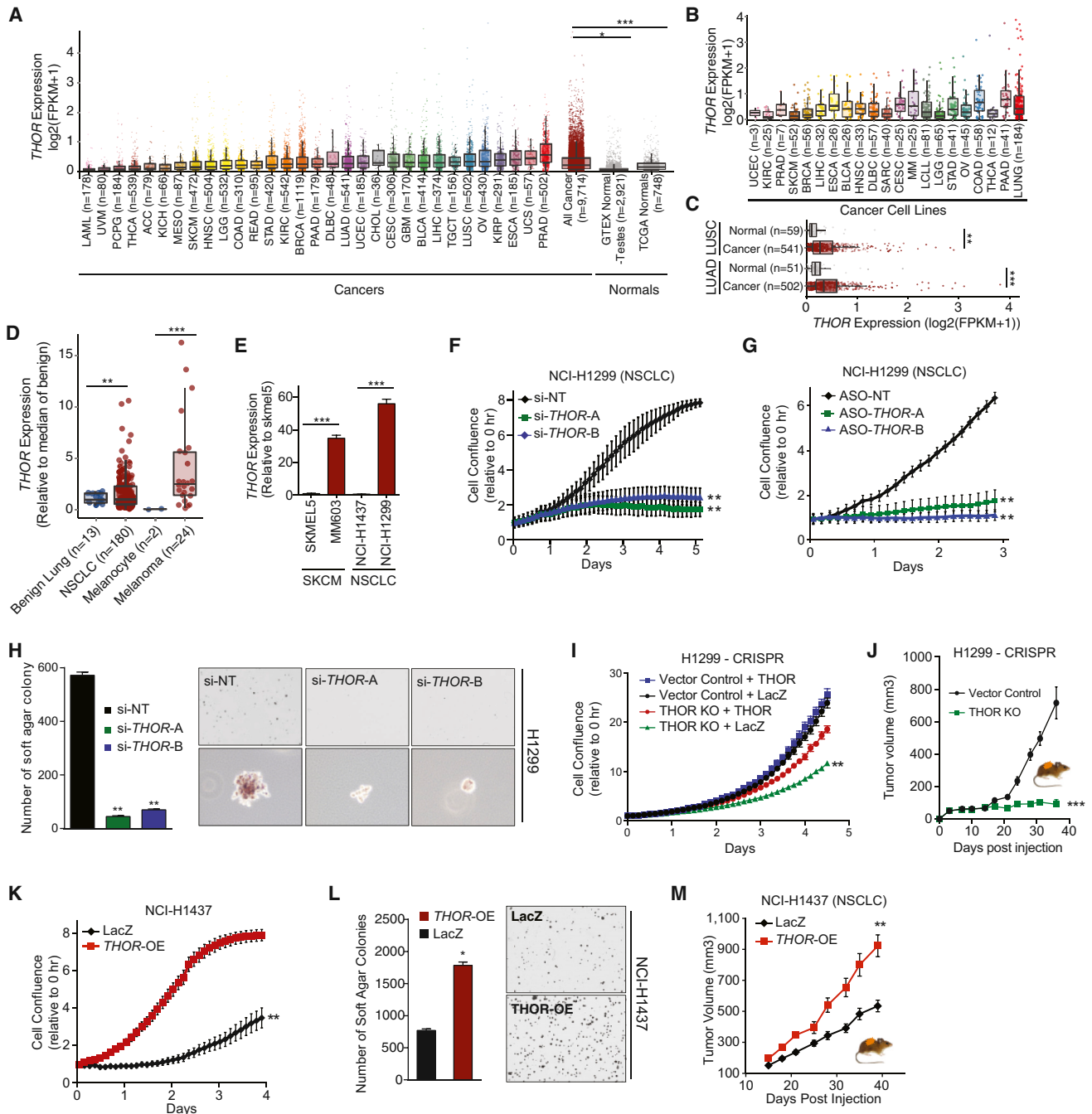
Figure 3. *THOR* Is Expressed in Cancers and Potentiates Tumorigenesis

(A) Expression of *THOR* in TCGA tumors and normal tissue samples from GTEx and TCGA.

(B) Expression of *THOR* in the CCLE cell line panel.

(C) Expression of *THOR* in the TCGA lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) samples represented alongside each tissue's matched normal samples.

(D) qRT-PCR validation in an independent tissue in LUAD (benign, n = 13; cancer, n = 180) and melanoma tissues (benign, n = 2; cancer, n = 24).

(E) Expression levels of *THOR* in two skin cutaneous melanoma (SKCM) and two non-small cell lung cancer (NSCLC) cell lines.

(F) Cell proliferation assays for NCI-H1299 cells treated with 2 independent *THOR* siRNAs.

(G) Cell proliferation assays for NCI-H1299 cells treated with 2 independent *THOR* ASOs.

(H) Anchorage-independent growth of H1299 cells transfected with non-targeting siRNA (si-NT) or two *THOR* siRNAs (siTHOR-A, siTHOR-B). (Left) Quantification of number of colonies. (Right) Representative image of surviving colonies and individual colony.

(I) Cell proliferation assay for NCI-H1299 cells with CRISPR-Cas9-mediated THOR knockout versus control in the context of LacZ and *THOR* overexpression.

(legend continued on next page)

Expression and Functional Implication of *THOR* in Human Cancers

In addition to its testis-specific expression in normal tissue, we identified expression of *THOR* in a myriad of cancer types, with little expression in non-testis normal tissue (Figure 3A) and in a number of cancer cell lines (Barretina et al., 2012), particularly in lung cancer cell lines (Figure 3B). Furthermore, *THOR* exhibited significant differential expression in lung adenocarcinoma (LUAD) and lung squamous carcinoma (LUSC) compared to normal lung samples (Figure 3C). This cancer-specific expression profile was corroborated in an independent cohort of tumor and normal lung in addition to melanoma and melanocytes obtained from the University of Michigan via qPCR (Figure 3D).

We performed knockdown of *THOR* via siRNA and antisense oligonucleotide (ASO) in H1299 and MM603, non-small-cell lung cancer (NSCLC) and melanoma cell lines with high levels of *THOR* (Figures 3E, S3A, and S3B) and observed a dramatic reduction in the proliferative capacity of these cells (Figures 3F, 3G, S3C, and S3D). siRNA and ASO knockdown of *THOR* in H1437 cells (lacking endogenous *THOR* expression; Figure 3E) exhibited no significant proliferation phenotype (Figures S3E and S3F). *THOR* knockdown resulted in reduced colony formation in soft agar via both ASO and siRNA knockdown (Figures 3H and S3G–S3I) and siRNA knockdown (Figure S3H).

Additionally, we generated a *THOR*-knockout cell line model via CRISPR-Cas9 technology with paired single-guide RNAs (sgRNAs) targeted to the conserved region of *THOR* transcript in the H1299 cell line. Multiple sgRNAs were utilized, targeting varying regions of *THOR* (Figure S3J), and a monoclonal population with robust knockout was selected for further use (Figures S3K and S3L). *THOR*-knockout H1299 cells exhibited significantly reduced cell proliferation (Figure 3I) with a recovery of phenotype with ectopic expression of *THOR* (Figure 3I). These results were also recapitulated in a mosaic population of knockout clones, suggesting that the monoclonal findings are not due to selection bias (Figures S3M–S3O). Further corroborating on-target effects, *THOR* mosaic knockout in H1437 cells (Figure S3P) did not result in reduced proliferation (Figure S3Q). Moreover, a mouse xenograft of H1299 cells containing *THOR* knockout exhibited markedly reduced tumor growth compared to control knockout cells (Figure 3J).

Cells with stable lentiviral *THOR* overexpression in H1437 and SKMEL5 (Figure S3R) exhibited significant increases in proliferative capacity (Figures 3K and S3S) and soft agar colony formation (Figures 3L and S3T). Additionally, murine tumor xenografts derived from cells stably overexpressing *THOR* in H1437 cells exhibited a significant proliferative advantage when compared to cells stably overexpressing LacZ control (Figure 3M). How-

ever, this finding was not significant in a murine xenograft using SMKEL5 cells (Figure S3U). Interrogation of the lentiviral plasmid via northern blotting revealed an unexpected long isoform of *THOR* in addition the isoform included in the plasmid (Figure S2A). 5′ and 3′ RACE identified a segment of plasmid expressed in the longer isoform (Figures S2J and S2K); however, *THOR*-targeting siRNAs did reduce levels of this isoform, suggesting functional fidelity of this longer isoform (Figure S2A).

Characterization of the *THOR*-IGF2BP1 Interaction

Cellular localization of *THOR* via qRT-PCR following cellular fractionation revealed expression in both the cytoplasm and nucleus (Figure S4A), a finding corroborated by single-molecule fluorescence ISH (Figures S4B–S4H). To further functionally characterize *THOR*, we identified potential *THOR* protein interactors via RNA pull-down followed by mass spectrometry (STAR Methods). Drawing from the sequence conservation of *THOR*, we leveraged multiple experimental conditions to identify conserved binding partners: (1) pull-down of h-*THOR* added to human H1299 cancer cell lysate, (2) pull-down of h-*THOR* added to zebrafish embryo lysate, (3) pull-down of z-*THOR* added to human H1299 cancer cell lysate, and (4) pull-down of z-*THOR* added to zebrafish embryo lysate (Table S3). In all conditions, pull-down of antisense *THOR* was utilized as a negative control. Two proteins, IGF2BP1 and IGF2BP3, were pulled down in all four conditions (Figure 4A). Interestingly, the IGF2BP proteins were also the only proteins pulled down by h-*THOR* in both the nuclear and cytoplasmic fractions of H1299 cells (Figure S4I), and IGF2BP1 has been reported to exhibit a cancer/testis expression pattern similar to that of *THOR* (Bell et al., 2013) (Figures S4J and S4K).

IGF2BP1, in conjunction with number of other proteins comprising the messenger ribonucleoprotein (mRNP) complex, has been implicated in mediating RNA stability and translation through its binding to a number of well-defined mRNA targets (Bell et al., 2013; Hafner et al., 2010; Weidensdorfer et al., 2009). Multiple members of the mRNP complex are pulled down by *THOR* in the various conditions (Figure 4A), and we confirmed that IGF2BP1, IGF2BP2, IGF2BP3, and YBX1 are present in the complex via immunoprecipitation and western blotting (Figure 4B). RNA immunoprecipitation (RIP) assays with antibodies against IGF2BP1-3, YBX1, STAU1, and HuR, followed by qRT-PCR for *THOR* and additional control RNAs, confirmed the specificity of the *THOR*-IGF2BP1 interaction (Figure 4C). Additionally, overexpression of *THOR* in H1437 produced a modest increase in the IGF2BP1 and IGF2BP3 interaction, suggesting a potential role for *THOR* in mediating the mRNP complex formation (Figure S4L).

(J) *THOR*-knockout NCI-H1299 cell line xenografts (n = 10) demonstrate decreased tumor growth relative to control samples (n = 10). Tumor volumes at each time point by caliper measurement are shown.

(K) Cell proliferation assay in NCI-H1437 cells stably transfected with *THOR* overexpression or LacZ control lentivirus. Data show mean ± SE from one of the two independent experiments.

(L) Anchorage-independent growth of LacZ or *THOR* overexpressing H1437 cells.

(M) *THOR*-overexpressing NCI-H1437 cell line xenografts (n = 10) demonstrate increased tumor growth relative to control LacZ samples (n = 10). Tumor volumes at each time point by caliper measurement are shown. Asterisk (*) indicates p ≤ 0.001 by a two-tailed Student's t test. Data show mean ± SEM from one of the two independent experiments.

For all panels, * indicates p ≤ 0.01, ** indicates p ≤ 0.001, and *** indicates p ≤ 0.0001 by a two-tailed Student's t test. Data show mean ± SE for all panels with error bars. See also Figure S3.
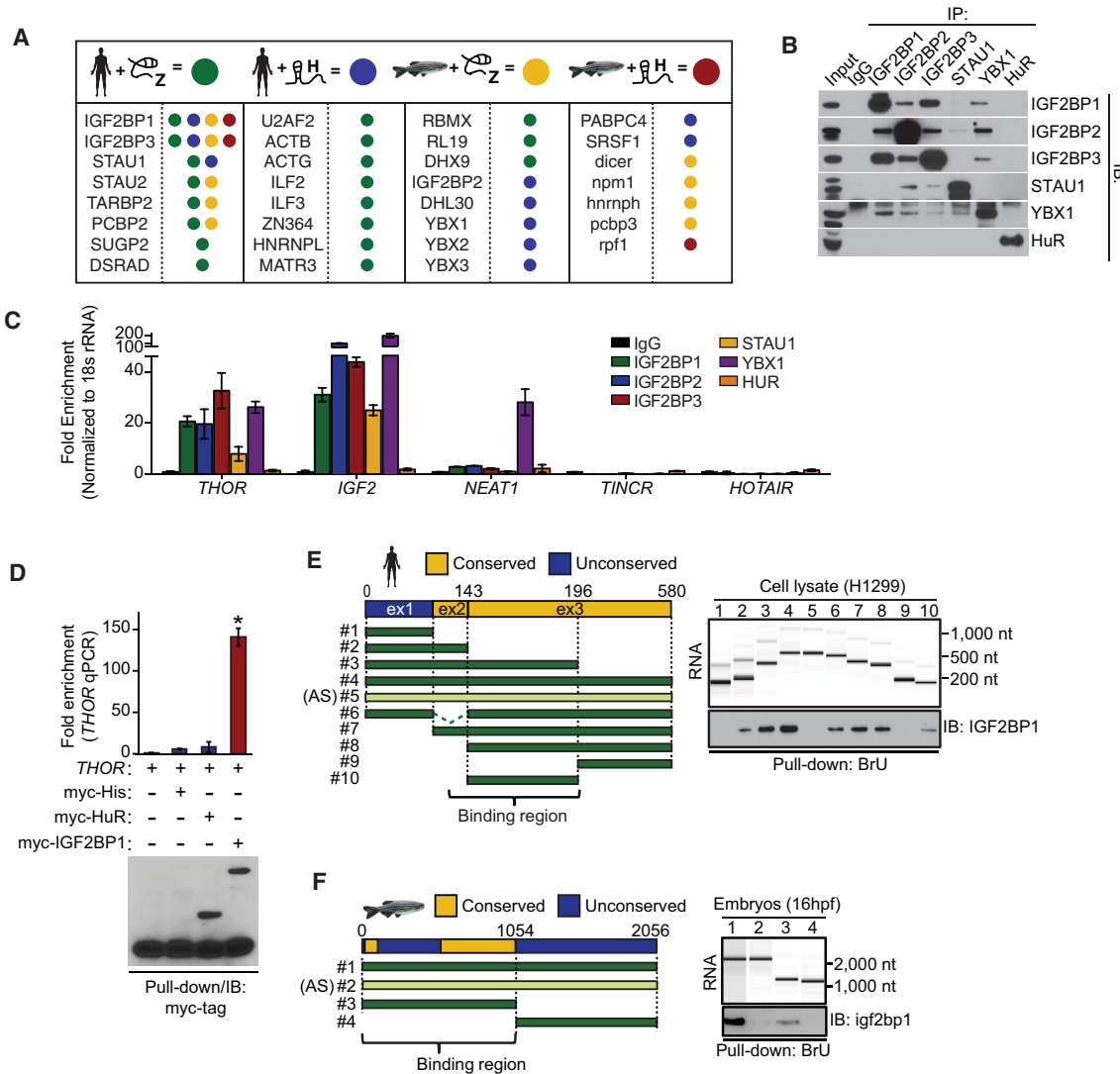
Figure 4. Conserved Interaction of *THOR* and IGF2BP1

(A) Table reporting the protein binding partners for *THOR* in four different experimental conditions of RNA pull-down analysis: zebrafish *THOR* added to human H1299 cell lysate (green), human *THOR* added to human H1299 cell lysate (blue), zebrafish *THOR* added to zebrafish embryo lysate (yellow), and human *THOR* added to zebrafish embryo lysate (red). All proteins bound in any condition are displayed in the table, and each dot represents binding in the respective condition.

(B) Immunoprecipitation-western-blotting analysis (IP-WB) for various components of the IGFBP complex and HuR as a negative control.

(C) qRT-PCR following RIP of IGF2BP1, IGF2BP2, IGF2BP3, STAU1, YBX1, HUR, and IgG in H1299 cells. Data show mean ± SD from one of the two independent experiments.

(D) *In vitro* RNA-protein binding assay. *In vitro* transcribed *THOR* added to purified myc-tagged proteins. *THOR* qRT-PCR was then performed following anti-myc pull-down.

(E) Schematic representation of human *THOR,* antisense-*THOR* (AS), and various deletion constructs generated to interrogate IGF2BP1 binding (left). Fragment sizes confirmed by PCR (right, top), and binding of each fragment to IGF2BP1 determined via pull-down of BRU-labeled RNA fragments (right, bottom) in H1299 cells.

(F) Schematic representation of zebrafish *THOR* constructs generated to study IGF2BP1 binding. Fragment sizes confirmed by PCR (right, top), and binding of each fragment to zebrafish igf2bp1 determined via pull-down of BRU-labeled RNA fragments (right, bottom) in 16-hpf embryos.

* indicates p ≤ 0.01 by a two-tailed Student's t test. Data show mean ± SE for all panels with error bars. See also Figure S4 and Table S3.

THOR and *IGF2* RNAs are pulled down by the various proteins of the mRNP complex, while negative control lncRNAs *NEAT1*, *TINCR*, and *HOTAIR* exhibit a substantially reduced extent of pull-down (Figure 4C). Nevertheless, HuR—a protein not in the mRNP complex (Figure 4B)—does not pull down these RNAs (Figure 4C), despite robust pull-down confirmed via immunoblot (Figure 4B). Direct binding interaction of *THOR* and IGF2BP1 was confirmed via addition of *in vitro* transcribed h-*THOR* to purified myc-tagged IGF2BP1 (Figure 4D).

Pull-down of various deletion isoforms of h-*THOR* followed by western blot for IGF2BP1 revealed that the region of *THOR* responsible for IGF2BP1 binding is in exon 2 and 3, the conserved region of h-*THOR* (Figure 4E). Additionally, this observation was also observed for z-*THOR*, wherein the 5′ conserved portion of z-*THOR* was sufficient to result in pull-down of the zebrafish igf2bp1 protein (Figure 4F). IGF2BP1 possesses 2 RNA-recognition-motif (RRM) domains and 4 Khomology (KH) domains (Figure S4M). Using multiple recombinant deletion isoforms of IGF2BP1, we revealed that deletion of the RRM domains did not affect *THOR* binding, while the KH1, KH3, and KH4 domains were found to be essential for THOR binding (Figure S4N).

THOR Regulates IGF2BP1's Target mRNAs Levels

IGF2BP1 has been reported to regulate the mRNA stability of a set of well-described target RNAs (Bell et al., 2013; Hämmerle et al., 2013) (*IGF2*, *CD44*, *KRAS*, *ACTB*, *PABPC1*, *GLI1*, *MYC*, *MAPT*, *CTNNB1*, *PPP1R9B*, *BTRC*, *PTEN*, and *H19*). Remarkably, the levels of nearly all IGF2BP1 targets were decreased upon knockdown of *THOR* in both H1299 and MM603 cells and conversely increased with stable overexpression of *THOR* in H1437 and SKMEL5 cells (Figures 5A and S5A). As expected, knockdown of *IGF2BP1* produced a similar reduction in its targets, while altering levels of *IGF2* and *CD44*—two of the canonical IGF2BP1 targets—failed to show a trend in the expression of IGF2BP1 targets (Figure 5A). CRISPR-mediated *THOR*-knockout H1299 cells displayed a similar reduction of expression of IGF2BP1 targets, with a reversal of phenotype when expressing ectopic *THOR* in these cells, suggesting a *trans* function for THOR (Figure S5B).

We hypothesized that the effects of *THOR* levels on IGF2BP1 targets (Figure 5A) may be explained by a *THOR*-mediated stabilization of the interaction of IGF2BP1 with its targets. Corroborating this hypothesis, qRT-PCR for IGF2BP1 targets *IGF2* and *CD44*, following IGF2BP1 RIP in the context of *THOR* knockdown, reduced IGF2BP1 binding of both *IGF2* and *CD44*, while overexpression of *THOR* increased *IGF2* and *CD44* binding. Knockdown of *IGF2* and *CD44* did not produce the same result (Figures 5B and S5C). Additionally, *THOR* overexpression substantially increased the mRNA stability of IGF2BP1 targets *IGF2* and *CD44* following actinomycin D treatment (Figure 5C) while having no effect on the stability of the *GAPDH* and *UBC*, two control mRNAs that do not interact with IGF2BP1 (Figure S5D). The *THOR* mRNA has a half-life of 14 hr (Figure S5E), which is longer than the dynamic range observed for the stabilization effects on *IGF2* and *CD44* (Figures 5C and S5D), confirming that *THOR* is present in cells long enough to exert these effects.

In addition to altering levels of *IGF2* (Figure 5A) and modulating *IGF2*-IGF2BP1 binding (Figure 5B), *THOR* is also successful in regulating the downstream signaling pathway of *IGF2* (Figures S5A and S5F), suggesting that the *THOR*-mediated changes in IGF2BP1 target expression are sufficient to result in functional downstream IGF2 signaling. Additionally, the proliferative advantage conferred by *THOR* overexpression was abrogated by reduction of *IGF2BP1* levels in those cells, with a particularly striking phenotype in the SKMEL5 cell line (Figure 5D), further

corroborating the functional relevance of the *THOR*-IGF2BP1 interaction. Additionally, knockdown of *IGF2BP1* in both H1299 and MM603 cells reduced cell proliferation (Figures S5G and S5H) and soft agar colony formation (Figures S5I and S5J). Overexpression of a deletion construct of *THOR* lacking the conserved IGF2BP1 binding sequence failed to exhibit the enhanced proliferation observed with full-length *THOR* overexpression (Figure 5E). These data suggest that the binding interaction between *THOR* and IGF2BP1 leads to functionally relevant consequences in cells that have implications for the oncogenicity of *THOR*.

To more broadly assess the transcriptional phenotype of *THOR* knockdown in comparison to *IGF2BP1* knockdown, differential expression was assessed for RNA-seq performed with two independent siRNAs targeted to *THOR*, *IGF2BP1*, and *HUR*. We observed significant overlap between the genes, with differential expression upon *THOR* and *IGF2BP1* knockdown (Figures 6A and 6B). Corroborating that the observed gene expression changes via siRNA knockdown are on-target effects, knockdown of *THOR* via two independent ASOs produced gene expression changes in line with those produced via siRNA knockdown (Figure S6A). Knockdown of *HUR*, however, did not recapitulate these gene-expression changes, serving as a negative control. The gene signatures most altered upon knockdown of *THOR* and *IGF2BP1* were also highly concordant (Subramanian et al., 2005) (Pearson r = 0.50; Figure 6C). The gene-signature changes following HUR knockdown, however, were not correlated to those following either *THOR* (Figure S6B) or *IGF2BP1* (Figure S6C) knockdown. Interestingly, two independent gene signatures associated with metastasis and relapse of melanoma (Kauffmann et al., 2008; Winnepenninckx et al., 2006) were among the top gene signatures altered upon knockdown of both *IGF2BP1* and *THOR* (Figure 6C), further implicating the *THOR*-IGF2BP1 relationship in cancer progression—particularly in melanoma.

To further implicate a functional relationship of *THOR* and IGF2BP1, we performed IGF2BP1 iCLIP (Huppertz et al., 2014). In H1437 cells overexpressing *THOR*, binding of IGF2BP1 on *THOR* was observed in the same region identified via deletion construct pull-downs (Figure 4E), while no binding was observed in the H1437 cells overexpressing the LacZ control (Figure S6D). 185 genes were identified as IGF2BP1 binding targets via iCLIP, and these genes were observed to have a significant increase in expression compared to genes that are not IGF2BP1 targets, as measured by RNA-seq (Figures S6E and S6F). Additionally, IGF2BP1 iCLIP in H1299 corroborated the localization of IGF2BP1 binding on *THOR* (Figure S6G). As above, the genes identified as binding targets of IGF2BP1 in the H1299 iCLIP experiment exhibited a significant reduction in expression upon *THOR* knockdown when compared to the genes without IGF2BP1 binding, although the magnitude of effect was lesser than in the H1437 experiment (Figures S6H and S6I). These data provide a high-throughput corroboration of the potential effects of THOR in mediating RNA stability of IGF2BP1 targets. Further leveraging these data, *cis* function for *THOR* was ruled out by the finding that genes near *THOR* do not exhibit gene-expression changes upon modulation of *THOR* levels (Figure S6J).
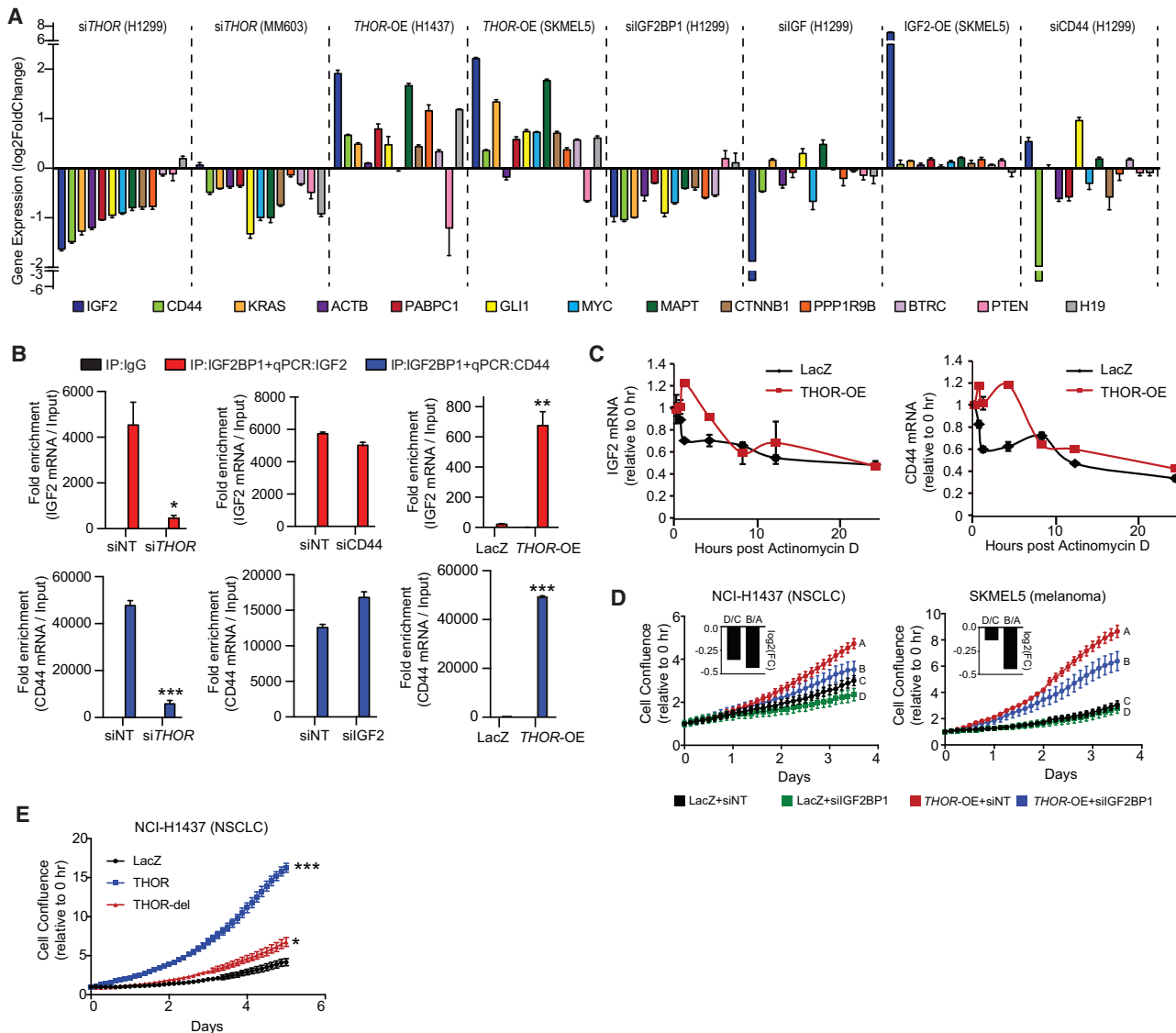
Figure 5. Interrogation of the Functional Relationship of *THOR* and IGF2BP1

(A) Bar plot depiction of the expression levels of 13 canonical IGF2BP1 target genes by qRT-PCR in various conditions: *THOR* knockdown, *THOR* overexpression, *IGF2BP1* knockdown, *IGF2* overexpression, and *CD44* knockdown. Data show mean ± SD from one of the two independent experiments.

(B) qRT-PCR expression levels for *IGF2* (red, black) and *CD44* (blue, black) following RIP of IGF2BP1 or IgG as negative control. RIPs performed in H1299 cells under various experimental conditions: *THOR* siRNA knockdown, *IGF2* siRNA knockdown, *CD44* siRNA knockdown, and *THOR* overexpression.

(C) qRT-PCR expression levels for *IGF2* (top) and *CD44* (bottom) following actinomycin D treatment in *THOR* or *LacZ* overexpressing H1347 cells.

(D) Cell proliferation assays for H1437 and SKMEL5 cells overexpressing *LacZ* control and *THOR* in the context of siRNA control and siRNA knockdown of *IGF2BP1*. Inset depicts the log2(fold change) comparing the proliferation at the final time point for control and *IGF2BP1* knockdown for the LacZ and *THOR*-overexpression settings.

(E) Cell proliferation assay in H1437 cells overexpressing full-length *THOR*, a *THOR* deletion mutant lacking the IGF2BP1 binding site, and LacZ control.

* indicates p ≤ 0.05 and ** indicates p ≤ 0.001 by two-tailed Student's t test. Data show mean ± SE for all panels with error bars. See also Figure S5.

THOR-Knockout Zebrafish Exhibit Fertilization Defects and Resistance to Melanoma Formation

Given that we observed sequence conservation of *THOR* (Figures 1A, 1B, 1D, and S1C–S1E), a conserved tissue expression pattern (Figures 2A, 2C, and 2D), and conservation of its binding interaction to IGF2BP1 (Figures 4A, 4E, and 4F), we set out to interrogate the potential function of *THOR* in a different animal model, extending the implications of its functionality beyond human cancer cell lines (Figures 3 and 4). The zebrafish animal model recently become a relevant model system for cancer investigation (Lieschke and Currie, 2007; White et al., 2013). Utilizing the CRISPR-Cas9 genome-editing system (Hwang et al., 2013; Sánchez-Rivera and Jacks, 2015), we produced a *THOR*-knockout zebrafish line. Two sgRNAs targeting the
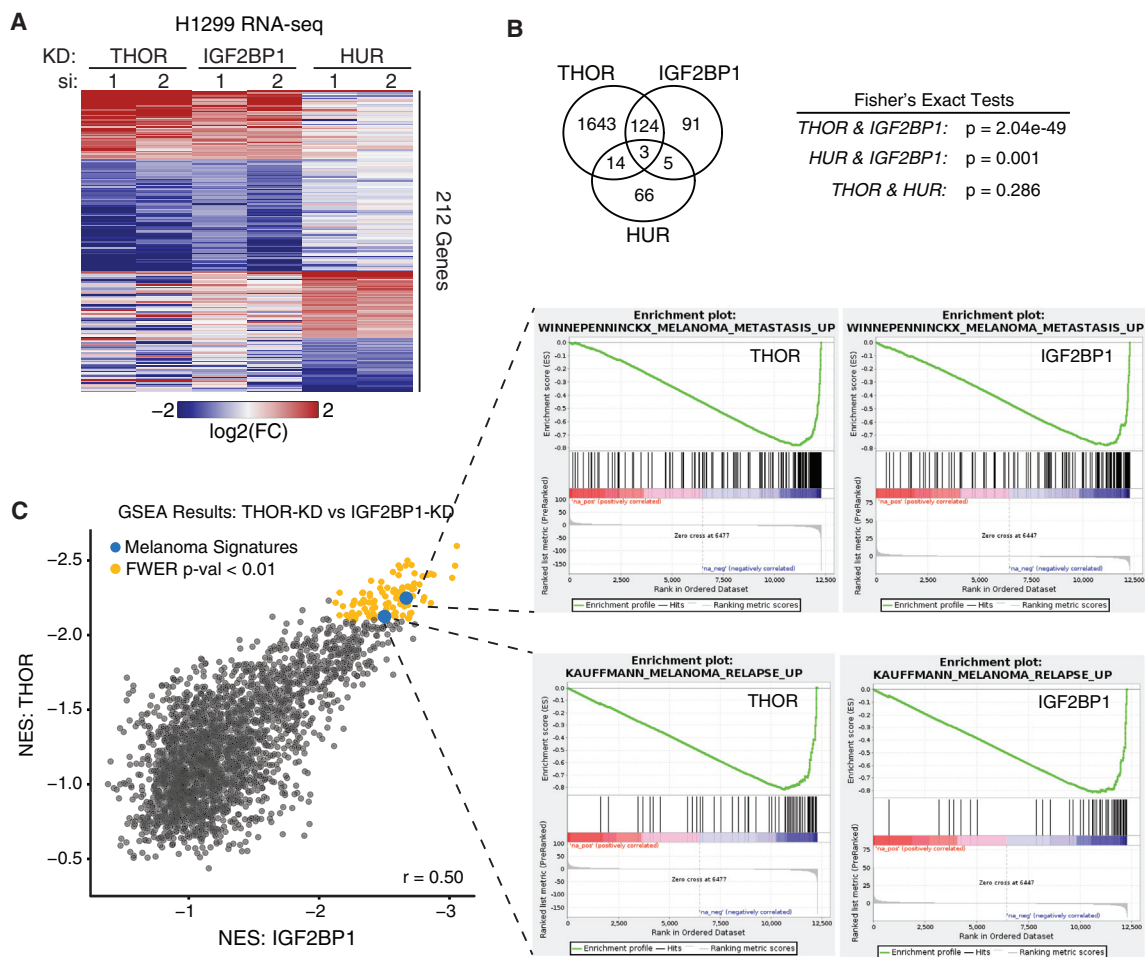
A

B

C

Figure 6. Shared Transcriptional Regulation by *THOR* and *IGF2BP1*

(A) Heatmap depicting the expression of the genes significantly differentially expressed (DESeq false discovery rate [FDR] < 0.05) in knockdown of *THOR* and *IGF2BP1* in H1299 cells in addition to those genes with significant differential expression in *HUR* knockdown. Expression depicted as the log2(fold change) for each siRNA compared to the non-targeting siRNA control.

(B) Venn-diagram depiction of the overlap for the significant differentially expressed genes in *THOR*, *IGF2BP1*, and *HUR* knockdown. Fisher's exact statistics shown on the right.

(C) Scatterplot depicting the gene set enrichment analysis (GSEA) performance for Molecular Signatures Database (MSigDB) v5.0 gene signatures with normalized enrichment score (NES) < 0 for both *THOR* and *IGF2BP1* knockdown (left). Signatures significant upon knockdown of both genes (family-wise error rate [FWER] p value < 0.01) depicted in gold. Two melanoma gene signatures depicted in blue. Pearson correlation coefficient shown in bottom right of scatterplot. GSEA plots for two significant melanoma signatures depicted for knockdown of *THOR* and *IGF2BP1* (right).

See also Figure S6.

conserved region of z-*THOR* (Figure S7A) and Cas9 mRNA were injected into zebrafish embryos, producing a mosaic F0 generation that was subsequently mated to wild-type zebrafish to generate heterozygous F1 offspring (*THOR*$^{+/-}$). The heterozygotes were then mated with one another to generate a population of homozygous *THOR*-knockout zebrafish (*THOR*$^{-/-}$) in the F2 generation (Figure 7A).

Upon generation of *THOR*$^{-/-}$ zebrafish, we observed a striking phenotypic effect on the fertility of *THOR*$^{-/-}$ zebrafish in comparison to wild-type zebrafish, with 55% of embryos from mating of *THOR*$^{-/-}$ zebrafish either dead or unfertilized 6 hr post fertilization (hpf) compared to only 11% from wild-type mating (Figure 7B). Moreover, when mating wild-type males to female

THOR$^{-/-}$ zebrafish, the fertilization defect was substantially diminished, while mating wild-type females to male *THOR*$^{-/-}$ zebrafish produced a significant fertilization defect in the zebrafish offspring (Figure 7C), supporting the role of *THOR* in the testis and suggesting a primary functional role for *THOR* in fertility. Moreover, expression of *THOR* was found to be isolated to spermatocytes in meiosis II at much higher levels than sperm at any other stage of development (Figure 7D), and testis of *THOR*$^{-/-}$ zebrafish contained fewer cells in meiosis II compared to wild-type zebrafish (Figures S7C and S7D). Of note, a preponderance of the most-altered gene signatures following *THOR* knockdown (Figures 6C, S6B, and S6C) were meiosis associated (Figures S7E and S7F). Within these signatures, we observed a
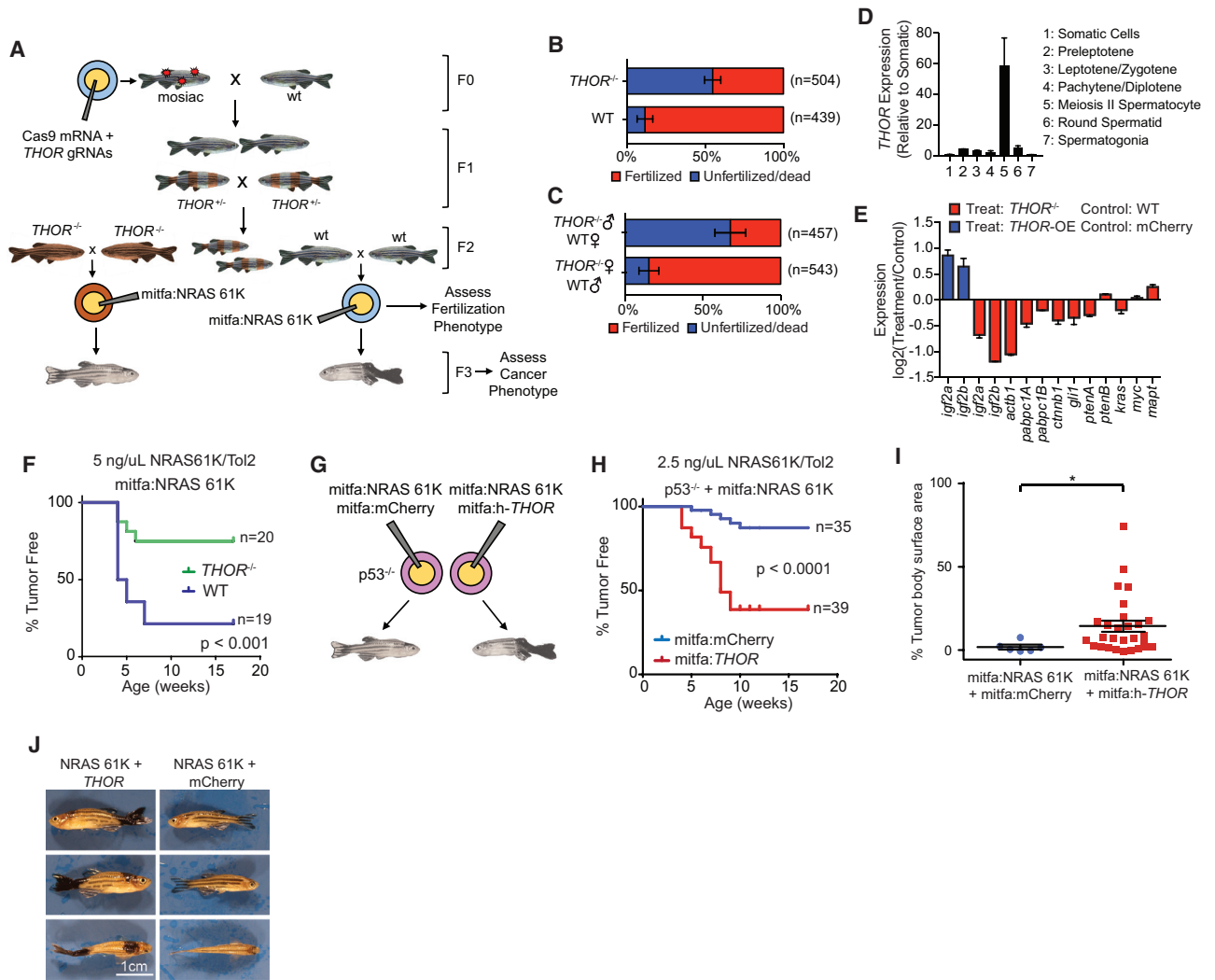
Figure 7. *THOR* Regulates Melanoma Onset in Zebrafish

(A) Schematic depicting creation of *THOR*-knockout zebrafish model.

(B) Fraction of fertilized zebrafish embryos derived from wild-type or *THOR*-knockout zebrafish (−/−) 6 hr following mating. * indicates $p \leq 0.01$ by a $\chi2$ test. Data show mean ± SD from two independent experiments.

(C) Fraction of fertilized 6-hpf zebrafish embryos derived from wild-type female crossed with *THOR*-knockout male or *THOR*-knockout female crossed with wild-type male. * indicates $p \leq 0.01$ by a $\chi2$ test. Data show mean ± SD from two independent experiments.

(D) Expression levels of z-*THOR* in testicular somatic cells as well as in 6 Hoechst-profiled subpopulations of testicular germ cells. Data show mean ± SD.

(E) Bar plot demonstrates expression of 12 zebrafish orthologs of the canonical IGF2BP1 target genes by qRT-PCR in zebrafish embryos. Expression represented as log2 of the fold change of either *THOR*[−/−] compared to wild-type embryos (red) or *THOR* overexpression compared to control mCherry. Data show mean ± SD from one of the two independent experiments.

(F) Kaplan-Meier curve of melanoma free period for mitfa promoter driven NRAS 61K zebrafish in either *THOR*[−/−] background or wild-type background. p values were determined using a log-rank test.

(G) Schematic describing the generation of the h-*THOR*-overexpression melanoma zebrafish model.

(H) Kaplan-Meier curve of melanoma-free period for p53[−/−] zebrafish co-injected with either mitfa promoter driven NRAS 61K + mitfa promoter driven human *THOR* or mitfa promoter driven NRAS 61K + mCherry. p values were determined using a log rank test.

(I) Percentage of body area covered in melanoma for mCherry and h-*THOR* injected zebrafish also containing mitfa-promoter-driven NRAS 61K. * indicates $p \leq 0.01$ obtained by a two-tailed Student's t test.

(J) Specimen example of NRAS 61K driven melanomas in zebrafish co-injected with mitfa-promoter-driven h-*THOR* or mitfa-promoter-driven mCherry in a p53[−/−] background.

Data show mean ± SE for all panels with error bars. See also Figure S7.

striking preponderance of upregulation in histone genes involved in meiosis. Many of these meiotic histone genes were identified as some of the most positively dysregulated genes upon *THOR* knockdown (Figure S7G). While the mechanism of the role of *THOR* in fertility needs further elucidation, we have shown evidence of a temperospatial localization of *THOR* within certain testis cells, particularly during meiosis II, and have shown corroborative evidence that those genes regulated by *THOR* are also implicated in meiosis.

Leveraging the zebrafish model to further investigate the role of *THOR* in oncogenesis, we interrogated gene expression of zebrafish embryos. Expression of *IGF2BP1* targets in *THOR*$^{-/-}$ zebrafish embryos was reduced compared to wild-type embryos, and expression of *igf2a* and *igf2b* was increased in zebrafish embryos ectopically overexpressed h-*THOR* (Figure 7E).

Additionally, to further interrogate *in vivo* function of *THOR*, we produced a zebrafish melanoma model that employs embryonic injection of human NRAS-K61 driven by the mitfa promoter (a zebrafish gene expressed in melanocytes), resulting in the formation of easily visible zebrafish melanoma (Figure 7A) (Dovey et al., 2009; Langenau et al., 2008) (for further details, see STAR Methods). Using this NRAS melanoma system, we observed a striking resistance to melanoma development in the *THOR*$^{-/-}$ zebrafish (Figure 7F). Of note, while the previous NRAS61K model of zebrafish melanoma required a p53 mutant background for tumorigenesis, we observed tumor growth in p53 wild-type zebrafish when we injected a high amount of NRAS61K/Tol2 into the single-cell embryo (5 ng/uL). Our ability to generate tumors in a p53 wild-type context is likely due to the F0 mosaicism and to the increased efficiency of transgenesis of the Tol2 system. The fertilization defect and resistance to melanoma mediated by *THOR* present compelling evidence for a conserved role of *THOR* in vertebrate physiology and pathophysiology.

Human *THOR* Enhances the Onset of Melanoma in Zebrafish

In order to further investigate the role of *THOR* in zebrafish melanoma development, we assessed the function of the addition of human *THOR* to zebrafish embryos utilizing the mitfa-promoter-driven injection method with injection of mCherry used as a negative control (Figure 7G) (Langenau et al., 2008). Injection was performed on p53 knockout (p53$^{-/-}$) zebrafish to enhance the melanoma phenotype observed as previously described (Dovey et al., 2009). In this model, embryos were injected with a lower concentration of NRAS61K/Tol2 (2.5ng/uL) than the *THOR*-knockout model (Figures 7A and 7F), resulting in a more indolent phenotype (Figure 7H), despite the loss of p53. Nevertheless, we show that loss of p53—in the context of *THOR* overexpression—significantly reduces tumor-free survival (Figure S7B). The overexpression of h-*THOR* in zebrafish was sufficient for a significant increase in not only the onset of melanoma development (Figure 7H), but also on the size of the melanoma tumors that developed (Figures 7I and 7J). NRAS 61K induced melanomas in both p53$^{-/-}$ and wild-type zebrafish (Figures 7J, S7I, and S7K) were positive for Melan-A (a mitf target gene and marker for melanoma in human specimens) staining, confirming that the lesions were in fact melanoma (Figures S7J and S7L). Thus, the striking ability of the human isoform of *THOR* to

promote melanoma in zebrafish argues for an evolutionarily conserved role of *THOR* in mediating cellular processes that are potentially involved in tumor development.

DISCUSSION

We have identified the first cancer/testis lncRNA, *THOR*, as a lncRNA with a cancer/testis expression pattern that exhibits a conserved interaction with IGF2BP1, potentially promoting oncogenesis. In doing so, we have defined a new class of lncRNA that may provide insight into our understanding of cancer biology, and in line with other cancer/testis antigens, this new class of lncRNA may provide potential for development of future cancer therapeutics (Sahu et al., 2015; Simpson et al., 2005). LncRNAs are highly abundant in the human transcriptome, but little is known regarding the ways in which most of these lncRNAs are functioning in cells. Thus, in this study, we aimed to define a new class of lncRNA and perform a thorough investigation of its functional relevance in cells. We therefore enforced significant evolutionary conservation as a selection criterion for nomination of *THOR* (Figures 1A and 1B) in order to leverage an animal model to robustly interrogate its role in cells. With a testis-specific tissue expression pattern observed in humans, mouse, and zebrafish, *THOR* presented the possibility of bearing an evolutionarily conserved function, strengthening its likelihood of performing an important role in many vertebrate species (Figures 2A, 2C, and 2D). Other interesting cancer/testis lncRNAs that are not highly conserved likely exist and present an exciting avenue for future studies. In line with our goal of identifying a conserved function for *THOR*, we identified protein-binding interactions for both human *THOR* and zebrafish *THOR* in human and zebrafish cells (Figure 4A). This method provided a powerful way to identify the *THOR*-IGF2BP1 interaction as one conserved across multiple vertebrate species. The binding of lncRNAs to RNA-binding proteins (RBPs) has been reported to be widespread (Li et al., 2015), and multiple efforts have been undertaken to characterize the functional implication of lncRNA-RBP interaction beyond promiscuous binding (Hasegawa et al., 2010; Lee et al., 2016; Li et al., 2014). Nevertheless, in this study, we show a conserved interaction of *THOR* with IGF2BP1 that has not been reported prior. While conservation of lncRNA sequence and function remains an active area of research, examples of conserved lncRNA-protein interactions such as *THOR* will help in furthering our functional understanding of lncRNAs.

Generation of the *THOR*-knockout cell lines, we were able to corroborate the findings of other RNAi methods, suggesting a robust functional phenotype mediated by *THOR*. The CRISPR-mediated knockout results are largely concordant with the RNAi findings, suggesting that much of the function of *THOR* is mediated by the RNA molecule itself and not necessarily its genomic location. Nevertheless, minor differences between the CRISPR knockout results and RNAi results may be evidence of an additional role of the *THOR* locus beyond the function of the transcript.

We subsequently implicated *THOR* in assisting IGF2BP1 in its regulation of multiple mRNA targets (Figure 5A), potentially via modulating the ability of IGF2BP1 to bind to these targets

(Figure 5B) and potentially regulate their transcriptional stability (Figures S6E–S6I). We also show that this *THOR*-IGF2BP1 relationship is associated with the proliferation cancer phenotype observed with *THOR* overexpression (Figures 5D and 5E). Post-transcriptional regulation of mRNAs is controlled tightly by their interaction with various RBPs (Bell et al., 2013; Decker and Parker, 2012; Glisovic et al., 2008). IGF2BPs are almost exclusively observed in the peri-nuclear region of cytoplasm, where they associate with target mRNAs in cytoplasmic mRNPs and play a central role in mRNA regulation (Bell et al., 2013). We have shown in this study that *THOR* binds to IGF2BP1 and regulates broad target mRNA stabilization. In light of the surprisingly long half-life of IGF2BP-RNA complexes *in vitro* (Nielsen et al., 2004), *THOR* may have a potential role in promoting the formation of "stable" protein-RNA complexes, although further experiments are required to determine the precise mechanism. It is worth noting that the IGF2BPs have been shown to largely localize to the cytoplasm (Bell et al., 2013). Thus, the cytoplasmic and nuclear localization of *THOR* suggests a potential function of *THOR* outside of its interaction with IGF2BP1.

While animal models have been employed to investigate lncRNAs (Sauvageau et al., 2013; Ulitsky et al., 2011), we present the first investigation of an oncogenic lncRNA employing a zebrafish cancer assay (Figure 7A). Moreover, many of these prior studies were limited to investigation of lncRNAs with little tissue expression in humans. *THOR*, however, is expressed in multiple species and is also widely expressed in a myriad of human cancers (Figure 3A), elevating its relevance to our understanding of normal and disease processes in humans. We show that loss of zebrafish *THOR* results in aberrant fertilization (Figures 7B and 7C), targeting this function to the male testis, and we also show that loss of *THOR* yields a resistance to melanoma formation (Figure 7F), implicating a function for z-*THOR* in the zebrafish. Interestingly, we also show that the human *THOR* is capable of utilizing zebrafish cellular machinery to also produce a striking cancer phenotype (Figures 7H and 7I). This trans-species function of *THOR* provokes interest, suggesting that its sequence and function have both been evolutionarily selected for. Investigating highly conserved lncRNAs like *THOR* elucidates key mechanisms through which lncRNAs function, and as the scientific community continues to investigate the multitude of unstudied novel lncRNAs, the zebrafish cancer model system presented here provides a powerful platform that can be used for further investigation of lncRNA function.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Murine Subcutaneous *In Vivo* Models
  - Zebrafish Maintenance
  - Zebrafish Melanoma Model
  - Cell lines
  - Tissue Samples

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quantitative Real-time PCR
  - Cell Proliferation Assays
  - *In Vivo* Xenograft Experiments
  - MS Data Analysis
  - Anchorage-independent soft agar colony formation Assay
  - RNA-seq Data Processing
  - RNA-Seq Differential Expression Testing
  - Gene signature GSEA analysis
- METHOD DETAILS
  - Conservation analysis
  - Coding potential assessment
  - RNA isolation and cDNA synthesis
  - RNA-ligase-mediated Rapid Amplification of cDNA Ends (RACE)
  - siRNA Knockdown Experiments
  - Overexpression Studies
  - Northern Blotting
  - Expression of recombinant protein
  - RNA *In Situ* Hybridization in Testis
  - Single-Molecule Fluorescence *in situ* Hybridization in Cell Lines
  - Western Blot Analysis
  - RNA Immunoprecipitation (RIP) Assay
  - RNA Pulldown Assay
  - Mass Spectrometry
  - RNA-protein Interaction Assay
  - RNA Endogenous Degradation Assay
  - iCLIP
  - iCLIP data analysis
  - Embryo GFP sorting
  - Cas9 Target Site Design, Vector Construction and *in vitro* RNA Synthesis
  - Generation of THOR knockout cell line
  - Microinjection of Zebrafish Embryos
  - DNA Isolation and PCR Analysis for Identifying Deletion
  - Germ cell sorting
  - Generation of THOR Knockout Zebrafish
  - Zebrafish Mosaic Melanoma Model
- DATA AND SOFTWARE AVAILABILITY
  - Data Resources

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

Bell, J.L., Wächter, K., Mühleck, B., Pazaitis, N., Köhn, M., Lederer, M., and Hüttelmaier, S. (2013). Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? Cell. Mol. Life Sci. *70*, 2657–2675.

Berghmans, S., Murphey, R.D., Wienholds, E., Neuberg, D., Kutok, J.L., Fletcher, C.D.M., Morris, J.P., Liu, T.X., Schulte-Merker, S., Kanki, J.P., et al. (2005). tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. Proc. Natl. Acad. Sci. USA *102*, 407–412.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. *25*, 1915–1927.

Calin, G.A., Liu, C.G., Ferracin, M., Hyslop, T., Spizzo, R., Sevignani, C., Fabbri, M., Cimmino, A., Lee, E.J., Wojcik, S.E., et al. (2007). Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell *12*, 215–229.

Consortium, T.G.; GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585.

Decker, C.J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb. Perspect. Biol. *4*, a012286.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dovey, M., White, R.M., and Zon, L.I. (2009). Oncogenic NRAS cooperates with p53 loss to generate melanoma in zebrafish. Zebrafish *6*, 397–404.

Fratta, E., Coral, S., Covre, A., Parisi, G., Colizzi, F., Danielli, R., Nicolay, H.J., Sigalotti, L., and Maio, M. (2011). The biology of cancer testis antigens: putative function, regulation and therapeutic potential. Mol. Oncol. *5*, 164–182.

Gaysinskaya, V., Soh, I.Y., van der Heijden, G.W., and Bortvin, A. (2014). Optimized flow cytometry isolation of murine spermatocytes. Cytometry A *85*, 556–565.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. *582*, 1977–1986.

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

Hämmerle, M., Gutschner, T., Uckelmann, H., Ozgur, S., Fiskin, E., Gross, M., Skawran, B., Geffers, R., Longerich, T., Breuhahn, K., et al. (2013). Posttranscriptional destabilization of the liver-specific long noncoding RNA HULC by the IGF2 mRNA-binding protein 1 (IGF2BP1). Hepatology *58*, 1703–1712.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. *22*, 1760–1774.

Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., and Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of xist RNA. Dev. Cell *19*, 469–476.

Holman, J.D., Tabb, D.L., and Mallick, P. (2014). Employing ProteoWizard to Convert Raw Mass Spectrometry Data. Curr. Protoc. Bioinformatics *46*, 1–9.

Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J., and Ule, J. (2014). iCLIP: protein-RNA interactions at nucleotide resolution. Methods *65*, 274–287.

Hudson, R.S., Yi, M., Volfovsky, N., Prueitt, R.L., Esposito, D., Volinia, S., Liu, C.-G., Schetter, A.J., Van Roosbroeck, K., Stephens, R.M., et al. (2013). Transcription signatures encoded by ultraconserved genomic regions in human prostate cancer. Mol. Cancer *12*, 13.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. *31*, 227–229.

Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner, J.R., Evans, J.R., Zhao, S., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. Nat. Genet. *47*, 199–208.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2014). Nucleic Acids Res. 42, D764–D770.

Kauffmann, A., Rosselli, F., Lazar, V., Winnepenninckx, V., Mansuet-Lupo, A., Dessen, P., van den Oord, J.J., Spatz, A., and Sarasin, A. (2008). High expression of DNA repair pathways is associated with metastasis in melanoma patients. Oncogene *27*, 565–573.

Kent, W.J. (2002). BLAT–the BLAST-like alignment tool. Genome Res. *12*, 656–664.

Khosravi-Far, R., White, M.A., Westwick, J.K., Solski, P.A., Chrzanowska-Wodnicka, M., Van Aelst, L., Wigler, M.H., and Der, C.J. (1996). Oncogenic Ras activation of Raf/mitogen-activated protein kinase-independent pathways is sufficient to cause tumorigenic transformation. Mol. Cell. Biol. *16*, 3923–3933.

Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L., and Gao, G. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. *35*, W345–W349.

Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant, J.M., Yost, H.J., Kanki, J.P., and Chien, C.B. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Dev. Dyn. *236*, 3088–3099.

Langenau, D.M., Keefe, M.D., Storer, N.Y., Jette, C.A., Smith, A.C., Ceol, C.J., Bourque, C., Look, A.T., and Zon, L.I. (2008). Co-injection strategies to modify radiation sensitivity and tumor initiation in transgenic Zebrafish. Oncogene 27, 4242–4248.

Lee, S., Kopp, F., Chang, T.C., Sataluri, A., Chen, B., Sivakumar, S., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. Cell 164, 69–80.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323.

Li, J.H., Liu, S., Zheng, L.L., Wu, J., Sun, W.J., Wang, Z.L., Zhou, H., Qu, L.H., and Yang, J.H. (2015). Discovery of Protein-lncRNA Interactions by Integrating Large-Scale CLIP-Seq and RNA-Seq Datsets. Front Bioeng. Biotechnol 2, 88.

Li, Z., Chao, T.C., Chang, K.Y., Lin, N., Patil, V.S., Shimizu, C., Head, S.R., Burns, J.C., and Rana, T.M. (2014). The long noncoding RNA THRIL regulates TNFα expression through its interaction with hnRNPL. Proc. Natl. Acad. Sci. U. S. A 111, 1002–1007.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740.

Lieschke, G.J., and Currie, P.D. (2007). Animal models of human disease: zebrafish swim into view. Nat. Rev. Genet. 8, 353–367.

Lin, M.F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics 27, i275–i282.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science 339, 823–826.

Mehra, R., Shi, Y., Udager, A.M., Prensner, J.R., Sahu, A., Iyer, M.K., Siddiqui, J., Cao, X., Wei, J., Jiang, H., et al. (2014). A novel RNA in situ hybridization assay for the long noncoding RNA SChLAP1 predicts poor clinical outcome after radical prostatectomy in clinically localized prostate cancer. Neoplasia 16, 1121–1127.

Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al.; GTEx Consortium (2015). Human genomics. The human transcriptome across tissues and individuals. Science 348, 660–665.

Nielsen, J., Kristensen, M.A., Willemoës, M., Nielsen, F.C., and Christiansen, J. (2004). Sequential dimerization of human zipcode-binding protein IMP1 on RNA: a cooperative mechanism providing RNP stability. Nucleic Acids Res. 32, 4368–4376.

Palanichamy, J.K., Tran, T.M., Howard, J.M., Contreras, J.R., Fernando, T.R., Sterne-Weiler, T., Katzman, S., Toloue, S., Yan, W., Basso, G., et al. (2016). RNA-binding protein IGF2BP3 targeting of oncogenic transcripts promotes hematopoietic progenitor proliferation. J. Clin. Invest 126, 1495–1511.

Pitchiaya, S., Androsavich, J.R., and Walter, N.G. (2012). Intracellular single molecule microscopy reveals two kinetically distinct pathways for microRNA assembly. EMBO Rep. 13, 709–715.

Poliakov, A., Russell, C.W., Ponnala, L., Hoops, H.J., Sun, Q., Douglas, A.E., and van Wijk, K.J. (2011). Large-scale label-free quantitative proteomics of the pea aphid-Buchnera symbiosis. Mol. Cell. Proteomics 10, M110.007039.

Prensner, J.R., and Chinnaiyan, A.M. (2011). The emergence of lncRNAs in cancer biology. Cancer Discov. 1, 391–407.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat. Methods 5, 877–879.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc 8, 2281–2308.

Sahu, A., Singhal, U., and Chinnaiyan, A.M. (2015). Long noncoding RNAs in cancer: from function to translation. Trends Cancer 1, 93–109.

Sánchez-Rivera, F.J., and Jacks, T. (2015). Applications of the CRISPR-Cas9 system in cancer biology. Nat. Rev. Cancer 15, 387–395.

Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., and Dobbs, D. (2007). Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. Nucleic Acids Res. 35, W599–W605.

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D. (2010). ZiFit (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res. 38, W462–W468.

Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M., et al. (2013). Multiple knockout mouse models reveal lincRNAs are required for life and brain development. eLife 2, e01749.

Scanlan, M.J., Gure, A.O., Jungbluth, A.A., Old, L.J., and Chen, Y.T. (2002). Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. Immunol. Rev. 188, 22–32.

Simpson, A.J.G., Caballero, O.L., Jungbluth, A., Chen, Y.-T., and Old, L.J. (2005). Cancer/testis antigens, gametogenesis and cancer. Nat. Rev. Cancer 5, 615–625.

St Laurent, G., Wahlestedt, C., and Kapranov, P. (2015). The Landscape of long noncoding RNA classification. Trends Genet. 31, 239–251.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550.

Suster, M.L., Kikuta, H., Urasaki, A., Asakawa, K., and Kawakami, K. (2009). Transgenesis in zebrafish with the Tol2 transposon system. Methods Mol. Biol. 561, 41–63.

Tapparel, C., Reymond, A., Girardet, C., Guillou, L., Lyle, R., Lamon, C., Hutter, P., and Antonarakis, S.E. (2003). The TPTE gene family: cellular expression, subcellular localization and alternative splicing. Gene 323, 189–199.

Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell 154, 26–46.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 147, 1537–1550.

Uren, P.J., Bahrami-Samani, E., Burns, S.C., Qiao, M., Karginov, F.V., Hodges, E., Hannon, G.J., Sanford, J.R., Penalva, L.O., and Smith, A.D. (2012). Site identification in high-throughput RNA-protein interaction data. Bioinformatics 28, 3013–3020.

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res. 41, e74.

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

Westerfield, M. (2007). The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio), Fifth Edition (Univ. Or. Press Eugene Book).

White, R., Rose, K., and Zon, L. (2013). Zebrafish cancer: the state of the art and the path forward. Nat. Rev. Cancer 13, 624–636.

Winnepenninckx, V., Lazar, V., Michiels, S., Dessen, P., Stas, M., Alonso, S.R., Avril, M.F., Ortiz Romero, P.L., Robert, T., Balacescu, O., et al.; Melanoma Group of the European Organization for Research and Treatment of Cancer (2006). Gene expression profiling of primary cutaneous melanoma and clinical outcome. J. Natl. Cancer Inst. 98, 472–482.

Zarnegar, B.J., Flynn, R.A., Shen, Y., Do, B.T., Chang, H.Y., and Khavari, P.A. (2016). irCLIP platform for efficient characterization of protein-RNA interactions. Nat. Methods 13, 489–492.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
IGF2	Sigma	SAB1408589; RRID: AB_10740650
CD44	Cell Signaling Technology	#3578; RRID: AB_2076463
IGF2BP1	MBL	RN007P; RRID: AB_1570640
HuR	Millipore	03-102; RRID: AB_10615425
Total H3	Cell Signaling Technology	#9715; RRID: AB_331563
Myc-Tag	MBL	M047-3; RRID: AB_591112
MEK	Cell Signaling Technology	#9146; RRID: AB_10694922
p-MEK	Cell Signaling Technology	#9121; RRID: AB_331648
ERK	Cell Signaling Technology	#9102; RRID: AB_330744
p-ERK	Cell Signaling Technology	#9101; RRID: AB_331646
Rabbit Polyclonal IgG	MBL	PM035; RRID: AB_10805234
IGF2BP2	MBL	RN008P; RRID: AB_1570641
IGF2BP3	MBL	RN009P; RRID: AB_1570642
STAU1	MBL	RN012P; RRID: AB_1570648
YBX1	MBL	RN015P; RRID: AB_1570651
Melan-A	DAKO	M719629-2; RRID: AB_2335691
Halo-tag	Promega	G9281; RRID: AB_713650
KRAS	Santa Cruz	sc-30; RRID: AB_627865
Bacterial and Virus Strains		
pLenti6/V5-DEST Gateway Vector	Invitrogen	V49610
One Shot Stbl3 Chemically Competent *E. coli*	Invitrogen	C7373-03
One Shot TOP10 Chemically Competent *E. coli*	Invitrogen	C404010
Biological Samples		
Lung cancer and paired non-tumoral lung tissues	University of Michigan Health System	N/A
Melanoma tissues	University of Michigan Hospitals Cutaneous Surgery and Oncology Program	N/A
Human normal tissues RNA, see Table S8	This paper	N/A
Mouse normal tissues RNA, see Table S8	This paper	N/A
Mouse embryos RNA	University of Michigan Transgenic Core	N/A
Zebrafish normal tissues RNA	This paper	N/A
Zebrafish embryos RNA	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
Actinomycin D	Sigma	A1410
Critical Commercial Assays		
GeneRacer Kit with SuperScript III RT and TOPO TA Cloning Kit	Invitrogen	L150201
NorthernMax-Gly Kit	Ambion	AM1946
EZ-Magna RIP RNA-Binding Protein Immunoprecipitation kit	Millipore	#17-701
RiboTrap Kit	MBL	RN1011/RN1012

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
RNA-seq and iCLIP raw data	This paper	SRA: PRJNA415317
Experimental Models: Cell Lines		
NCI-H1299	ATCC	CRL-5803
NCI-H1437	ATCC	CRL-5872
MM603	Sigma	11072801-1VL
SK-MEL-5	ATCC	HTB70
Experimental Models: Organisms/Strains		
Mouse: CB17SCID	Charles River	Strain code: 236
Zebrafish: AB* (incross of AB and local wild type strain)	This paper	N/A
Zebrafish: tp53(zdf1/+)	Berghmans et al., 2005	RRID: ZFIN_ZDB-GENO-151014-5
Oligonucleotides		
Primers, see Table S4	This paper	N/A
siRNA, see Table S5	This paper	N/A
FISH probes, see Table S6	This paper	N/A
Recombinant DNA		
IGF2BP1 cDNA	Gegecopoea	NM_006546.3
IGF2 cDNA	Origene	SC119786
MLM3613	Addgene	#42251
DR274	Addgene	#42250
pSpCas9(BB)-2A-GFP (PX458)	Addgene	#48138
pBabe NRAS 61K	Addgene	#12543
pDESTtol2pA5	Tol2/Gateway kit	#684
p5E-MCS	Tol2/Gateway kit	#228
pME-MCS	Tol2/Gateway kit	#237
pME-mCherry	Tol2/Gateway kit	#386
p3E-polyA	Tol2/Gateway kit	#302
Software and Algorithms		
ZiFiT	Sander et al., 2007, 2010	http://zifit.partners.org/ZiFiT/
CRISPR Design	Zhang Lab, MIT 2015	http://crispr.mit.edu
Kallisto	Bray et al., 2016	https://pachterlab.github.io/kallisto/download
DESeq2	Anders and Huber, 2010	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
GSEA	Subramanian et al., 2005	http://software.broadinstitute.org/gsea/doc/GSEAUserGuideFrame.html
CPAT	Wang et al., 2013	http://rna-cpat.sourceforge.net
PhyloCSF	Lin et al., 2011	https://github.com/mlin/PhyloCSF/wiki
CPC	Kong et al., 2007	http://cpc.cbi.pku.edu.cn
MSConvert	Holman et al., 2014	http://proteowizard.sourceforge.net/tools.shtml
MASCOT	Matrix Science	http://www.matrixscience.com/search_intro.html
RSEM	Li and Dewey, 2011	https://deweylab.github.io/RSEM/
FastX Toolkit	Hannon Lab, CSHL	http://hannonlab.cshl.edu/fastx_toolkit
STAR	Dobin et al., 2013	https://github.com/alexdobin/STAR
Piranha	Uren et al., 2012	http://smithlabresearch.org/software/piranha/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Arul M. Chinnaiyan (arul@med.umich.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Murine Subcutaneous *In Vivo* Models

All experimental procedures were approved by the University of Michigan Institutional Animal Care and Use Committee (IACUC). Male CB17 SCID mice (Charles River Laboratories) aged 5-7 weeks were used for all studies. Animals were housed in a pathogen-free environment with 12-hour light-dark cycle and free access to food and water. Xenografts were generated by subcutaneous injection of cells into the bilateral posterior dorsal flanks (n = 10 xenografts per cell line). Xenografts were measured weekly using a digital caliper, with endpoint established as a volume of 1000 mm^3. Upon reaching this endpoint, or exhibiting signs of end-stage illness, mice were euthanized and xenografts resected. Resected specimens were divided in half, with portions allocated, respectively, for fixation in 10% buffered formalin and snap freezing. For THOR overexpression experiments, xenografts were established with injection of 0.25×10^6 NCI-H1437 LacZ or THOR overexpressing (THOR-OE) cells in a Matrigel scaffold (BD Matrigel Matrix, BD Biosciences). For THOR CRISPR knockout experiments, 1×10^6 vector control or THOR knockout NCI-H1299 cells were injected. For melanoma xenograft experiments, 1×10^6 LacZ or THOR overexpressing SK-MEL-5 cells were injected.

Zebrafish Maintenance

Zebrafish (*Danio rerio*) were maintained at 28-29°C under a 14 h light: 10 h dark cycle in a filtered freshwater recirculation system and fed three times daily as described previously (Westerfield, 2007). All breedings were conducted after approximately 90 days post fertilization, when fish were sexually mature. Previously separated male and female fish were introduced into a breeding tank and eggs were collected. Embryos were raised at 28.5°C and staged in hours post fertilization (hpf) according to standard procedures. The AB* strain, which was initially obtained from crosses between the AB strain (Zebrafish International Resource Center, ZIRC) and local wild-type strain and subsequently inbred in the lab, was used for generation of transgenic lines and mutants described in this paper. The *tp53*$^{zdf1/+}$ was obtained from ZIRC (Berghmans et al., 2005).Approval for zebrafish research was obtained from the University Committee on the Use and Care of Animals (UCUCA) of the University of Michigan.

Zebrafish Melanoma Model

In previous zebrafish melanoma models, naked DNA containing NRAS61K was injected into single cell zebrafish embryos containing a mutant p53 background (Dovey et al., 2009). The naked DNA transgenesis model is largely inefficient and requires mating of transgenic zebrafish to the F1 generation for successful generation of the transgenic fish, which can negatively impact transgenesis that can be deleterious to animal viability (Suster et al., 2009). Here, we employ a Tol2 integration system which has been shown to produce markedly more robust transgenesis (Kwan et al., 2007; Suster et al., 2009). In addition to the more robust integration of the NRAS transgene, this system enables generation of a mosaic F0 generation that can be utilized in functional experimentation. The time frame of transcription and translation of the transposase injected into the embryos is longer than the time to replication of the single cell embryo, resulting in mosaic expression of the transgene in the adult zebrafish. This phenomenon circumvents the selection against particularly stressful transgenic events in the naked DNA model, enabling robust expression of NRAS61K in these Tol2 mediated transgenic zebrafish.

Cell lines

NCI-H1299, NCI-H1437 and SK-MEL-5were obtained from the American Type Culture Collection (Manassas, VA). MM603 was from Sigma. Cell lines were maintained using standard media and conditions. Specifically, NCI-H1299, NCI-H1437 and MM603 were maintained in RPMI 1640 (Invitrogen) plus 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. SK-MEL-5 was maintained in EMEM (Invitrogen) plus 10% FBS and 1% penicillin-streptomycin. All cell lines were grown at 37°C in a 5% CO$_2$ cell culture incubator and genotyped for identity at the University of Michigan Sequencing Core and tested routinely for Mycoplasma contamination. *THOR* or control-expressing cell lines were generated by cloning *THOR* or control into the pLenti6 vector (Invitrogen) using pCR8 non-directional Gateway cloning (Invitrogen) as an initial cloning vector and shuttling to plenti6 using LR clonase II (Invitrogen) according to the manufacturer's instructions. Stably-transfected NCI-H1437 and SK-MEL-5 cells were selected using blasticidin (Invitrogen). All lentiviruses were generated by the University of Michigan Vector Core.

Tissue Samples

All human subject studies were approved by the Institutional Review Board of the University of Michigan. The lung cancer and paired non-tumoral lung tissues were obtained from patients undergoing curative cancer surgery during the period from 1991 to 2012 at the University of Michigan Health System. None of the patients included in this study received any preoperative radiation or chemotherapy. All melanoma tissues were procured from the University of Michigan Hospitals Cutaneous Surgery and Oncology Program with appropriate informed consent. Resected specimens were frozen in liquid nitrogen and then stored at −80°C until use. Total RNA

panels from human and mouse normal tissues were purchased from Clontech and Zyagen. Mouse embryos were obtained from the University of Michigan Transgenic Core. Zebrafish tissues and embryos were obtained from AB strain wild-type zebrafish.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of genes were performed via qPCR and RNA-seq. Mass spectrometry was utilized to quantify protein abundance. Quantitation of cell proliferation, soft agar colony formation, and mouse xenograft growth were performed on multiple cell lines.

In relevant figures, Figure Legends denote the level of statistical significance for the test utilized in the various tests. Asterisks define degree of significance as described in the Figure Legends.

Quantitative Real-time PCR

Quantitative Real-time PCR (qRT-PCR) was performed using Power SYBR Green Mastermix (Applied Biosystems, Foster City, CA) on an Applied Biosystems 7900HT Real-Time PCR System. All oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA) and are listed in Table S4. The housekeeping genes, *GAPDH*, *HMBS* and *UBC*, were amplified as controls. Fold changes were calculated relative to housekeeping genes and normalized to the median value of the lung benign samples.

Cell Proliferation Assays

Proliferation experiments were carried out by plating $1\text{-}2 \times 10^4$ cells or $0.15\text{-}0.3 \times 10^4$ cells in 24-well or 96-well plates respectively and grown in regular media. Growth rate was monitored by IncuCyte live-cell imaging system (Essen Biosciences) for the specified durations.

In Vivo Xenograft Experiments

All experimental procedures were approved by the University of Michigan Committee for the Use and Care of Animals (UCUCA). Male mice (CB17SCID) aged 5-7 weeks were injected with 0.25×10^6 NCI-H1437 LacZ or THOR overexpressing (THOR-OE) cells with a Matrigel scaffold (BD Matrigel Matrix, BD Biosciences) in the posterior dorsal flank region (n = 10 per cell line). For THOR CRISPR knockout experiment, 1×10^6 vector control or THOR knockout H1299 cells were injected in the dorsal flank region of CB17SCID mice (n = 10 per cell line). For the melanoma xenograft experiment, 1×10^6 LacZ or THOR overexpressing SKMEL5 were injected subcutaneously into CB17SCID mice (n = 10 per cell line). In all murine xenograft experiments, tumor measurement was taken twice weekly using a digital caliper.

MS Data Analysis

The resulting spectrum files were transformed into MGF format by MSConvert software (Holman et al., 2014) and interrogated by MASCOT 2.4 search engine using human UniProt database version 15 concatenated with reverse sequences for estimation of false discovery rate (FDR) and with a list of common contaminants (40729 entries in total). The search parameters were as follows: full tryptic search, 2 allowed missed cleavages, peptide charges +2 and +3 only, MS tolerance 1 Da, MS/MS tolerance 0.5 Da. Permanent post-translational modifications was: cysteine carbamidomethylation. Variable post-translational modifications were: protein N-terminal acetylation, Met oxidation and N-terminal Glutamine to pyro-Glutamate conversion. The remaining analysis was performed as previously described (Poliakov et al., 2011). To summarize, the minimal ion score threshold was chosen such that a peptide false discovery rate (FDR) below 1% was achieved. The peptide FDR was calculated as: $2 \times$ (decoy_hits)/(target + decoy hits). Spectral counts for all detected proteins were assembled using an in-house written Python script. The adjustment of spectral counts was done by the same script as in (Poliakov et al., 2011). All raw spectral counts are described in Table S3.

Anchorage-independent soft agar colony formation Assay

For H1437 and H1299 soft-agar colony formation assay, $1\text{-}3 \times 10^3$ cells were suspended in DMEM containing 0.3% agar, 10% fetal bovine serum, and layered on DMEM containing 0.6% agar, 10% FBS in 6-well plate. After 2 weeks incubation, colonies were stained with iodonitrotetrozolium chloride (Sigma) for overnight. Visible colonies were enumerated from two replicate wells.

RNA-seq Data Processing

RNA-sequencing reads were quantified to the human transcriptome (GENCODEv25) using Kallisto (v0.43.0) (Bray et al., 2016). GENCODEv25 GTF was obtained from GENCODE (Harrow et al., 2012), and transcriptome fasta file was produced using the *rsem-prepare-reference* function of RSEM (version 1.2.26) (Li and Dewey, 2011). Kallisto index was generated using the *kallisto index* function. Transcript level quantification obtained using the *kallisto quant* function. Gene level expression obtained by summing the TPM values for all transcripts within each gene.

RNA-Seq Differential Expression Testing

Differentially expressed genes were obtained by comparing non-targeting shRNA control to each of the two replicates for the three genes tested (i.e., *THOR*, *IGF2BP1,* and *HUR*) using DESeq2 (Anders and Huber, 2010). Significantly differentially expressed genes were defined as genes with a greater than 20.75 log-fold-change with a q-value < 0.05.

Gene signature GSEA analysis

For each gene a rank list was generated by ordering each gene in the differential expression analysis by the DESeq2 (Anders and Huber, 2010) log-fold-change value (*log2foldchange)* by the q-value (*padj).* These rank lists were used in a weighted, pre-ranked GSEA (Subramanian et al., 2005) analysis against MSigDBv5 (Liberzon et al., 2011). Significant associations were determined for any gene set having an FWER p value below 0.01.

METHOD DETAILS

Conservation analysis

Evolutionary conservation of transcripts was assessed via the fraction of significantly conserved bases (p $\leq$ 0.01, phyloP algorithm), and the most conserved 200nt sliding window (phastCons scores averaged within each window). For contiguous sliding window conservation an average PhastCons probability of 0.9986 was used to identify ultraconserved elements as previously described (Iyer et al., 2015). PhyloP and phastCons scores were obtained from the UCSC Genome Browser (Karolchik et al., 2014).

Coding potential assessment

Coding potential for *THOR* was assessed using the CPAT (Wang et al., 2013) tool, PhyloCSF (Lin et al., 2011), and CPC (Kong et al., 2007) tool. CPAT and PhyloCSF were run using the command line tools. For PhyloCSF, the multiz alignment for 46 vertebrate species for the sequence conservation of *THOR* and *MYC* was obtained using the conservation track from the UCSC genome browser for GRCh38 (Karolchik et al., 2014). The CPC tool was run using their online tool (http://cpc.cbi.pku.edu.cn). Ribosomal profiling data was obtained using the GWIPS-viz genome browser (http://gwips.ucc.ie).

RNA isolation and cDNA synthesis

Total RNA from human and mouse normal tissues were purchased from Clontech and Zyagen. Mouse embryos were obtained from the University of Michigan Transgenic Core. Zebrafish tissues or embryos were obtained from AB strain wild-type fish. Nuclear and cytoplasmic fractions were separated using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo) according to the manufacturer's instructions. Total RNA was isolated using miRNeasy Mini Kit (QIAGEN) with DNase I (QIAGEN) digestion according to the manufacturer's instructions. RNA integrity was verified on an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). cDNAs weresynthesized from total RNA using Superscript III (Invitrogen) and random primers (Invitrogen).

RNA-ligase-mediated Rapid Amplification of cDNA Ends (RACE)

5′ and 3′ RACE was performed using the GeneRacer RLM-RACE kit (Invitrogen) according to the manufacturer's instructions. RACE PCR products were obtained using Platinum Taq High Fidelity polymerase (Invitrogen), the supplied GeneRacer primers, and appropriate gene-specific primers indicated in Table S4. RACE-PCR products were separated on a 2% agarose gels, bands excised and the extracted DNA (Gel Extraction kit, QIAGEN) were cloned into pCR4-TOPO vector (Invitrogen), and sequenced bidirectionally using M13 forward and reverse primers at the University of Michigan Sequencing Core. At least four colonies were sequenced for every gel product that was purified and the data was analyzed using Sequencher software (GeneCodes).

siRNA Knockdown Experiments

Knockdown experiments were carried out in approximately 1 - 2 × 10^5 cells plated in 100 mm dishes. While *THOR* knockdown in MM603 cells was achieved with two sequential transfections (at 24 hr and 48 hr post-plating) with 50 μM experimental siRNA oligos or non-targeting controls, for *THOR* knockdown in H1299 cells only one siRNA transfection (24hr post plating) was done. Only one transfection was performed for all protein coding gene knockdowns. Knockdowns were performed with RNAiMAX (Invitrogen) in OptiMEM media and its efficiency was determined by qRT-PCR 96 hr post-plating. All siRNAs were purchased from Dharmacon and their sequences (in sense format) are listed in Table S5.

Overexpression Studies

The *THOR* expression construct were generated by amplifying the full-length transcript from NCI-H1299 cells and subcloning into the pLenti6 expression vector (Invitrogen), LacZ constructs were used as controls. Following Sanger sequencing (University of Michigan Sequencing Core) confirmation of the inserts, lentiviruses were generated at the University of Michigan Vector Core. NCI-H1437 and SK-MEL-5 cells were infected with lentiviruses expressing *THOR* or LacZ and stable pools and clones were generated by blasticidin selection (Invitrogen). The THOR deletion constructs were also generated by amplifying by PCR using the full-length transcript as a template and were subcloned into the pLenti6 expression vector (Invitrogen).

Northern Blotting

Northern blotting was performed using the NorthernMax-Gly Kit (Ambion) following the manufacturer's protocol. Briefly, 20-30 ug of total RNA was denatured with Glyoxal loading dye solution for 30 minutes at 50°C, and separated on a 1% agarose glyoxal gel. The RNA was then transferred to Nylon Membrane (Roche) by capillary blotting with the transfer buffer and cross-linked with UV light (UV Stratalinker 1800). The membrane was subjected to a prehybridization step by incubation in Ultrahyb buffer (Ambion) at 68°C

for 1 hour. The membrane was incubated at 68°C overnight with antisense p32 labeled RNA probe in UltraHyb buffer. Following washing in accordance with the NorthernMax-Gly kit protocol, the membranes were exposed to HyBlot CL autoradiography film (Denville Scientific). The primer sequences used for generating the probes are given in Table S4.

Expression of recombinant protein

IGF2BP1 cDNA (NM_006546.3) was purchased from Gegecopoea. IGF2BP1 coding region was amplified by PCR and cloned into pFN19A (HaloTag7) T7 SP6 Flexi vector (Promega). IGF2BP1 deletion constructs were generated by inverse PCR using primers described in Table S3. All clones were verified by DNA sequencing. The HaloTag fusion proteins were synthesized by incubating 3 μg plasmid with *in vitro* TNT Quick-coupled Transcription/Translation System (Promega). Synthesized proteins were subjected to RNA Pulldown Assay.

RNA *In Situ* Hybridization in Testis

THOR ISH was performed on thin (approximately 4 μm thick) tissue sections (Advanced Cell Diagnostics, Hayward, CA), as described previously (Mehra et al., 2014). Appropriate batch positive and negative controls demonstrated expected staining patterns (data not shown). Slides were examined for THOR ISH signals in morphologically intact cells and scored manually by a study pathologist (R.M.). Specific THOR ISH signal was identified as brown, punctate dots.

Single-Molecule Fluorescence *in situ* Hybridization in Cell Lines

smFISH was performed as described (Raj et al., 2008), with some minor modifications. Cells were grown on 8-well chambered cover-glasses, formaldehyde fixed and permeablized overnight at 4°C using 70% ethanol. Cells were rehydrated in a solution containing 10% formamide and 2x SSC for 5 minutes and then treated with 10nM FISH probes for 16 h in 2x SSC containing 10% dextran sulfate, 2 mM vanadyl-ribonucleoside complex, 0.02% RNase-free BSA, 1 μg/μL E.coli tRNA and 10% formamide at 37°C. After hybridization the cells were washed twice for 30 minutes at 37°C using a wash buffer (10% formamide in 2x SSC). Cells were then mounted in solution containing 10 mM Tris/HCl pH 7.5, 2x SSC, 2 mM trolox, 50 μM protocatechiuc acid (PCA) and 50nM protocatechuate dehydrogenase (PCD). FISH samples were imaged in 3 dimensions using HILO illumination as described (Pitchiaya et al., 2012). Images were processed using custom-written macros in ImageJ. Analysis routines comprised of 3 major steps: background subtraction, Laplacian of Gaussian (LoG) filtering and thresholding. Spots with intensity above set threshold are represented in images. All probes were obtained from Biosearch technologies and are listed in Table S6.

Western Blot Analysis

Western blot analysis was performed according to standard procedures using Immobilon-P filters (Millipore) and an Enhanced Chemiluminescence detection system (GE Healthcare). Details of the primary antibodies used are listed in Table S7.

RNA Immunoprecipitation (RIP) Assay

RIP assays were performed using a Millipore EZ-Magna RIP RNA-Binding Protein Immunoprecipitation kit (Millipore, #17-701) according to the manufacturer's instructions. RIP-PCR was performed using total RNA as input controls and 1:150 of RIP RNA product was used per PCR reaction. The antibodies (3 - 5 ug of antibody per RIP reaction) used for RIP are described in Table S7.

RNA Pulldown Assay

RNA-pull down assays were performed using a RiboTrap Kit (MBL, RN1011/RN1012) according to the manufacturer's instructions. Briefly, 5-bromo-UTP (BrU) was randomly incorporated into the *THOR* RNA upon transcription using *THOR* full-length or deleted fragments PCR products as templates. Next Anti-BrdU antibodies conjugated with protein G beads (Invitrogen), were bound to the *in vitro* synthesized RNA before incubating with NCI-H1299 cell lysates for 4 hr. Finally, the samples were washed, eluted, and subjected to Mass spectrometry analysis.

Mass Spectrometry

The samples were treated with SDS-PAGE loading buffer supplied with 10 mM DTT for 5 min at 85°C. The proteins were alkylated by the addition of iodoacetamide to the final concentration of 15 mM. The samples were subjected to SDS-PAGE and the whole lanes were cut out and digested with trypsin in-gel for 2 hours. The resulting peptides were extracted, dried and resuspended in 0.1% formic acid with 5% acetonitrile prior to loading onto a trap EASY-column (Thermo Scientific) coupled to an in-house made nano HPLC column (20 cm x 75 um) packed with LUNA C18 media. Analysis was performed on Velos Pro mass spectrometer (Thermo Scientific) operated in data-dependent mode using 90-min gradients in EASY-LC system (Proxeon) with 95% water, 5% acetonitrile (ACN), 0.1% formic acid (FA) (solvent A), and 95% ACN, 5% water, 0.1% FA (solvent B) at a flow rate of 220 nl/min. The acquisition cycle consisted of a survey MS scan in the normal mode followed by twelve data-dependent MS/MS scans acquired in the rapid mode. Dynamic exclusion was used with the following parameters: exclusion size 500, repeat count 1, repeat duration 10 s, exclusion time 45 s. Target value was set at 104 for tandem MS scan. The precursor isolation window was set at 2 m/z. The complete analysis comprised two independent biological replicates.

RNA-protein Interaction Assay

The *in vitro* transcribed BrU labeled RNA were heated at 92°C for 2 min (to remove secondary structure), and incubated with recombinant myc-tagged proteins in RIP buffer (150 mM KCl, 25 mM Tris pH 7.4, 0.5 mM DTT, 0.5% NP40, 1 mM PMSF and protease inhibitor (Roche Complete Protease Inhibitor Cocktail Tablets) for 3chr at 4°C. RNA–protein complexes of interest were then partially purified with anti-myc magnetic beads (Thermo) and the products were treated with proteinase K, to remove the protein components leaving the RNAs intact. The recovered RNAs were extracted using miRNeasy Mini Kit as described above.

RNA Endogenous Degradation Assay

Cells were treated with 5 μg/ml Actinomycin D (Sigma) and collected in Quiazol at the indicated time points after treatment. Purified RNA was subjected cDNA synthesis and qRT-PCR as described above. The slopes for decay plots were determined by simple linear regression, and transcript half-life was calculated as the x intercept at y = 0.5, using GraphPad Prism.

iCLIP

iCLIP was performed as previously described (Huppertz et al., 2014; Palanichamy et al., 2016). Briefly, H1299, H1437-LacZ and H1437-THOR cells were cross-linked with UV light (UV Stratalinker 1800). After cell lysis, RNA was partially digested using RNase I (Life Technologies, AM2295), and IGF2BP1-RNA complexes were immunoprecipitated with anti-IGF2BP1 antibody (MBL International Corporation) immobilized on protein A–coated magnetic beads (Invitrogen). After 3′ end dephosphorylation by T4 PNK (NEB, M0201L), RNAs were ligated at their 3′ ends to a 3′ Preadenylated RNA adaptor, radioactively labeled by p32-γ, and run in MOPS-based protein gel electrophoresis. After transferring to a nitrocellulose membrane, protein-RNA complexes 15–80 kDa above free protein were cut from the membrane. We used SDS based RNA recovery platform as described previously (Zarnegar et al., 2016). We used the reverse transcription primers containing a 6-nt experiment-specific barcode within an 8-nt random barcode at their 5′ end to mark individual cDNA molecules. cDNA were size purified in TBE gel, circularized by CircLigase II (Cambio, CL9025K), annealed to an oligonucleotide complementary to the cleaved site and cut using BamHI (New England Biolabs). Linearized cDNAs were then PCR-amplified using AccuPrine SuperMix I (Invitrogen, 12342-010) and subjected to high throughput sequencing using Illumina HiSeq.

iCLIP data analysis

PCR duplicates were initially removed by collapsing identical reads. The iCLIP reads contained 8 random bases before the barcode, serving to distinguish reads arising from PCR amplification from reads arising from multiple RNA species. iCLIP eeads were first filtered for sequencing quality using the fastq_quality_filter tool in the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit) with the "-Q33 25" and "-p 80" flags. The fastx_collapser tool was used to collapse duplicate reads with the "-Q33" flag. Barcodes were trimmed from reads using the fastx_clipper tool, and random bases were trimmed using the fastx_trimmer tool also from the FASTX-Toolkit package.

Trimmed and deduplicated reads were then mapped to the GRCh38 genome using STAR (Dobin et al., 2013) using the "EndToEnd" option for the "—alignEndsType" flag, and "0.08" for the "—outFilterMismatchNoverLmax" flag. RT-stops were identified as the 5′ base in aligned reads, and a custom BED file was created for a window of 15 bases up and downstream of the RT stop. These 30BP windows surrounding the RT-stops were then used to identify peaks using Piranha (Uren et al., 2012) with the following commands: "-b 30 –s –p 0.01." Genes were identified as having IGF2BP1 binding if they were identified to have an exonic Piranha peak for both iCLIP replicates from the H1437 cells overexpressing *THOR*.

Embryo GFP sorting

48 hpf embryos were harvested and dechlorinated with Pronase (2 mg/ml) in E2 medium (15 mM NaCl, 0.5 mM KCL, 2.7 mM CaCl2, 1 mM MgSO4, 0.7 mM NaHCO3, 0.15 mM KH2PO4, 0.05 mM Na2HPO4). After deyolking by pipetting with 200 ul tip in 1/2 Ginsberg Fish Ringer without Calucium (55 mM NaCl, 1.8 mM KCL, 1.25 mM NaHCO3), embryos were re-suspended in the Protease Medium (0.25% Tripsin, 1mM EDTA in PBS pH = 8.0) and incubated for 40 min at 28 C with homogenizing with 200 ul tip every 10 min. After adding 100 ul FBS to stop reaction, cells were centrifuged for 3 min at 3,000 rpm, washed by Suspension Medium (08 mM CaCl2, 1% FBS in Leibovitz medium L-15 (GIBCO, 21083-027)) once, and filtered through strainer (352235, Falcon). GFP positive cells were sorted and collected in Quiazol followed by RNA extraction. Cell sorting and data analysis were performed by University of Michigan Flow Cytometry Core using MoFlo Astrios cell sorter (Beckman Coulter).

Cas9 Target Site Design, Vector Construction and *in vitro* RNA Synthesis

The plasmids MLM3613 bacterial Cas9 expression vector ((Addgene plasmid # 42251), (Mali et al., 2013)) and DR274 sgRNA expression vector ((Addgene plasmid # 42251), (Mali et al., 2013)) were purchased from Addgene (Cambridge, MA). We selected two sgRNA targets with ZIFIT Targeter (http://zifit.partners.org/zifit/Introduction.aspx) to generate a deletion of the conserved portion of THOR in zebrafish. For each target we annealed the oligonucleotide pairs and ligated into Bsa I-linearized DR274. sgRNAs were transcribed from Dra I-linearized templates using the MEGAscript T7 Kit (Ambion). Cas9 mRNA was transcribed *in vitro* with the mMESSAGE mMACHINE T7 ULTRA kit (Ambion). RNAs were purified by RNA Clean & Concentrator (Zymo Research) and re-dissolved in RNase-free water.

Generation of THOR knockout cell line

The plasmids pSpCas9(BB)-2A-GFP (PX458) (Addgene plasmid # 48138) (Ran et al., 2013) was purchased from Addgene (Cambridge, MA). We selected two sgRNA targets with CRISPR Design (http://crispr.mit.edu) to generate a deletion of the conserved portion of human THOR. For each target, we annealed the oligonucleotide pairs and ligated into BbsI-linearized PX458 plasmid. Cells were transfected with two vectors using Lipofectamine 3000 (Life Technologies) according to the manufacturer's instructions. 48 hours post-transfection, mosaic cells were genotyped and subjected to further experiments. To obtain monoclonal clones, GFP-positive cells were FACS sorted as a single cell into 96-well plate. After culturing for 3 weeks, cells are distributed into two 24 well plates followed by PCR-based genotyping. A clone showing deletion of the targeted region in THOR was used for further analysis. Single-cell sorted cells obtained after transfection of the empty PX458 construct was used as a negative control. Cell sorting and data analysis were performed by University of Michigan Flow Cytometry Core using MoFlo Astrios cell sorter (Beckman Coulter).

Microinjection of Zebrafish Embryos

One-cell stage embryos were microinjected with 250 ng/ul Cas9 mRNA and 150 ng/ul of each sgRNAs by using a pneumatic pico-pump (PV-820, World Precision Instrument).

DNA Isolation and PCR Analysis for Identifying Deletion

For embryonic gDNA extraction, 20 pooled embryos were lysed in 20 μL lysis buffer (10mM Tris HCl pH8.0, 2mM EDTA, and 0.2% Triton) containing proteinase K (10 μg/mL) at 55°C for 2 hr followed by 95°C for 10 minutes. 1 uL of lysate was used directly for genotyping PCR performed for40 cycles of 10 s at 95°C, 30 s at 60°C, and 60 s at 72°C after initial denaturing for 30 s at 95°C. PCR products were analyzed by 2% agarose gel. All genotyping primers are listed in Table S4.

Germ cell sorting

Germ cell sorting was performed as described previously (Gaysinskaya et al., 2014). Briefly, zebrafish testes were placed in 6 mL Collagenase I/Dnase I solution (200 U/ml Collagenase type I (Sigma-Aldrich) and 5 μg/ml DNase I (Invitrogen) in Gey's Balanced Salt Solution (GBSS) (Sigma-Aldrich)) and shaken at 150 rpm for 10 min at 35°C. The temperature and agitation speed were the same for all subsequent incubation steps. The testes were gently pipetted halfway into the 10 minute incubation. Tubules were settled for 2 min at room temperature (RT), then the supernatant, enriched in interstitial testicular cells (somatic cells), was harvested. 6 mL Collagenase I/Dnase I/Trypsin solution (200 U/ml Collagenase type I, 5 μg/ml DNase I and 0.025% Trypsin (GIBCO) in GBSS) was added to the pellet and the tubules were gently pipetted. Halfway into the 25 minute digestion period, 60 μL of 2.5% Trypsin was added, and the tubules were pipetted again. At the end of the incubation time, pipetting was repeated. The resulting cell-dense suspension was passed through a Nylon cell strainer (Falcon). To the resulting filtered cell suspension 10 μL of 1 mg/ml DNase I and 10 μL of 10 mg/ml Hoechst 33342 (Life Technologies) were added and incubated for 20 min. Halfway into the 20 minute period, the suspension was pipetted. At the end of incubation, 600 μL of FBS was added to inactivate the trypsin. After determining the cell number, the suspension was spiked with 10 μL of 1 mg/ml DNase I, and stained with Hoechst dye for the final 6 μg Hoechst/million cells. The suspension was incubated for 25 min. The cells were then stained with 10 μL of PI (Sigma-Aldrich) at RT. Cell sorting and data analysis were performed by University of Michigan Flow Cytometry Core using MoFlo Astrios cell sorter (Beckman Coulter). Hoechst was excited using 375 nm laser, and the dye's wide emission spectrum detected in two distinct channels: the "Ho Blue" (450/40 nm band-pass filter) and the "Ho Red" (670 nm long pass filter). Cells from each subpopulation were sorted and subjected to qRT-PCR.

Generation of THOR Knockout Zebrafish

F0 zebrafish was crossed to wild-type AB* to generate F1 embryos that were screened for THOR deletion. F0 zebrafish that were able to produce germ-line deletion of THOR were crossed to produce F1 heterozygotes, which were subsequently genotyped and crossed to generate THOR homozygous. THOR homozygotes and matched wild-type fish were used for phenotypic analyses.

Zebrafish Mosaic Melanoma Model

All transgenic constructs were made using the Tol2/Gateway kit (a gift from Dr. Kristen Kwan) (Kwan et al., 2007). Full-length human NRAS 61K was amplified from pBabe NRAS 61K construct (a gift from Channing Der (Addgene plasmid # 12543) [Khosravi-Far et al., 1996]) subcloned into BglII and BamHI restriction enzyme sites. GFP was amplified and subcloned into SalI and BamHI sites of pME entry vector. THOR transcript was amplified from THOR expression plasmid and cloned into SalI and BamHI sites of pME entry vector. The mitfa promoter was amplified using gDNA extracted from embryos as a template, and cloned into p5'E entry vector; and polyA tail was cloned into p3'E entry vector. These were assembled into the Tol2 destination vector using MultiSite Gateway Technology system (Invitrogen). 2.5 ng/μL of mitfa:NRAS 61K was co-injected into one-cell stage of p53−/− embryos(Berghmans et al., 2005) with mitfa: THOR or mitfa:mCherry (25 ng/ul each) with 2.5 ng/ul of Tol2 mRNA. For injections into THOR −/− embryos and their corresponding wild-type embryos, 5 ng/μL of mitfa:NRAS 61K was injected into one-cell embryos with 5 ng/ul of Tol2 mRNA. Zebrafish were inspected weekly, then euthanized when 17 weeks old and fixed in 4% paraformaldehyde overnight. After taking photos, they were then decalcified in 0.5 M ethylenediaminetetraacetic acid before paraffin embedding and sectioning.

Staining and immunohistochemistry were done using the standard techniques by the University of Michigan URAM Core. Percentage of melanoma area per body was calculated using ImageJ software. Antibodies used for immunohistochemistry are described in Table S7.

DATA AND SOFTWARE AVAILABILITY

Data Resources
All RNA-seq and iCLIP have been deposited in SRA. The accession number for these experiments reported in this paper is SRA: PRJNA415317. All questions related to the zebrafish data should be addressed to Weibin Zhou at weibin.zhou.duke.edu.

Figure S1. Conservation of Protein-Coding Genes and Genomic Organization of the Long Non-coding RNA *THOR*, Related to Figure 1

(A) Scatterplot depicting the distribution of basewise transcript conservation levels (x axis) and the average conservation for the best 200bp window (y axis) for all protein coding genes expressed at 1 FPKM or more in the top 1% of TCGA samples. Full transcript conservation levels were measured using the fraction of

conserved bases (PhyloP p < 0.01). Sliding window conservation levels were measured using the average PhastCons score across 200bp regions along the transcript. Green points indicate transcripts with 200bp windows that meet the criteria for 'ultraconserved' regions.

(B) Expression of lncRNA CRNDE among the GTEx normal tissue RNA-seq dataset, spanning a myriad of different normal tissue types. C-E, UCSC genomic browser view of *THOR* represented in the UCSC browser for (C) human GRCh37, (D) mouse GRCm38, and (E) zebrafish Zv9. *THOR* structure depicted along with H3K4me3 histone marks (ENCODE), conservation (Phylop and PhastCons) and Multiz 100 vertebrate alignment.
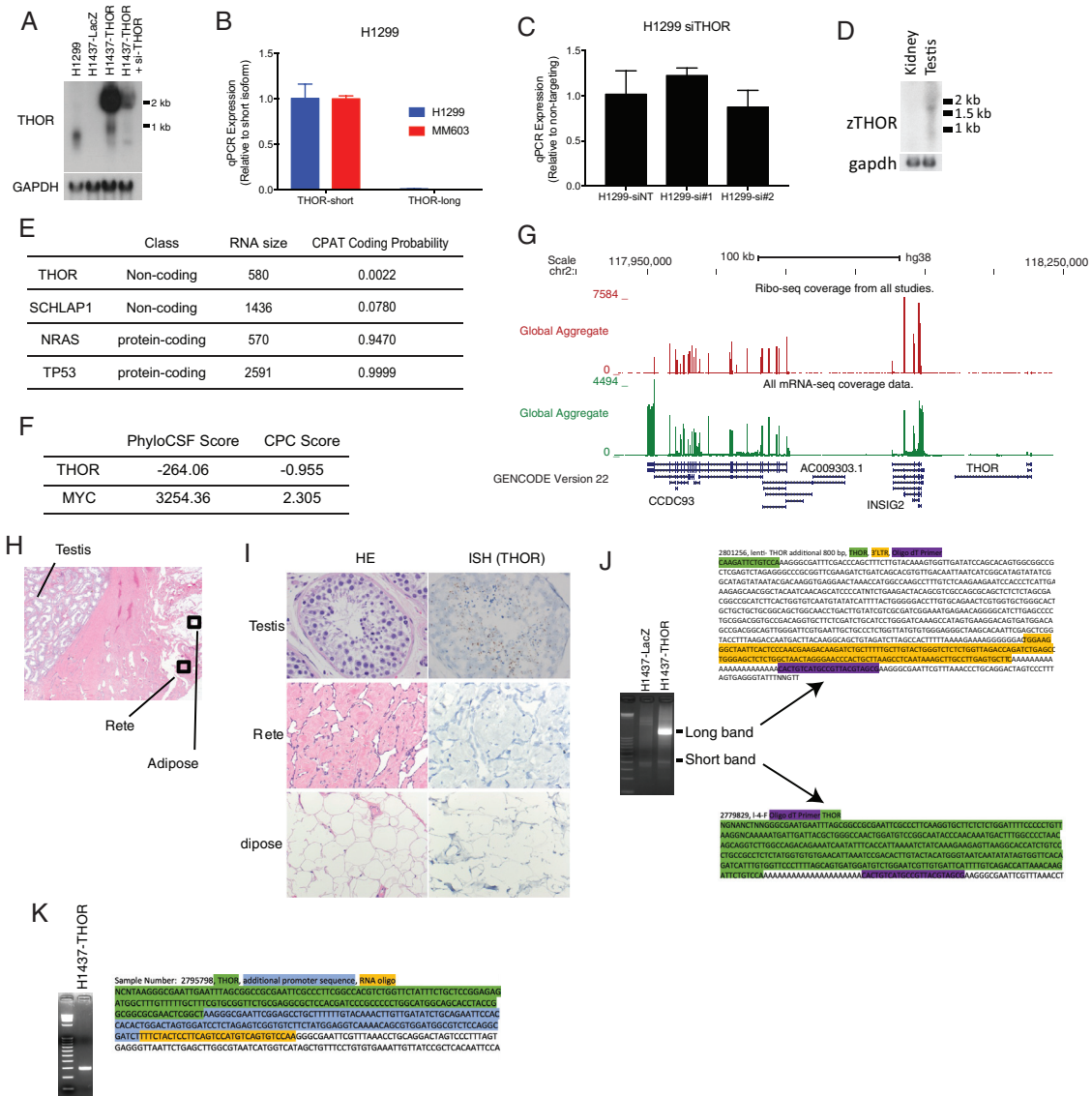
Figure S2. Characterization of *THOR* Transcript, Coding Potential, and Tissue Expression, Related to Figure 1

(A) Northern blot of endogenous *THOR* in H1299 cells, and of H1437 cells expressing LacZ control, *THOR*, and *THOR* with the addition of siRNA targeting *THOR*. Blot of gapdh provided as a control.

(B) Bar plot depicting the qPCR expression of the long versus short THOR isoform.

(C) qPCR expression of the long THOR isoform following addition of siRNA.

(D) Northern blot of *THOR* in zebrafish kidney and testis. Blot of GAPDH provided as a control.

(E) 5′ RACE for the THOR transcripts expressed by the lentiviral system. PCR agarose gel (left) confirms single band used in Sanger sequencing (right).

(F) 3′ RACE for the THOR transcripts expressed by the lentiviral system. PCR agarose gel (left) shows two bands utilized in Sanger sequencing (right).

(G) Coding probability scores for the transcripts were assessed by Coding Potential Assessment Tool (CPAT). *NRAS* and *TP53* used as positive control, and *SCHLAP1* as a negative control.

(H) Coding probability scores for the PhyloCSF and CPC tools for *THOR* and *MYC*. Values less than 0 suggest a lack of coding potential.

(I) Genome browser depiction of the *THOR* locus with aggregate ribosomal profiling track (red), aggregate poly-A RNA-seq track (green) and GENCODE v22 genome annotation obtained from the GWIPS-viz ribo-seq genome browser.

(J) H&E image of the testis and surrounding tissue architecture.

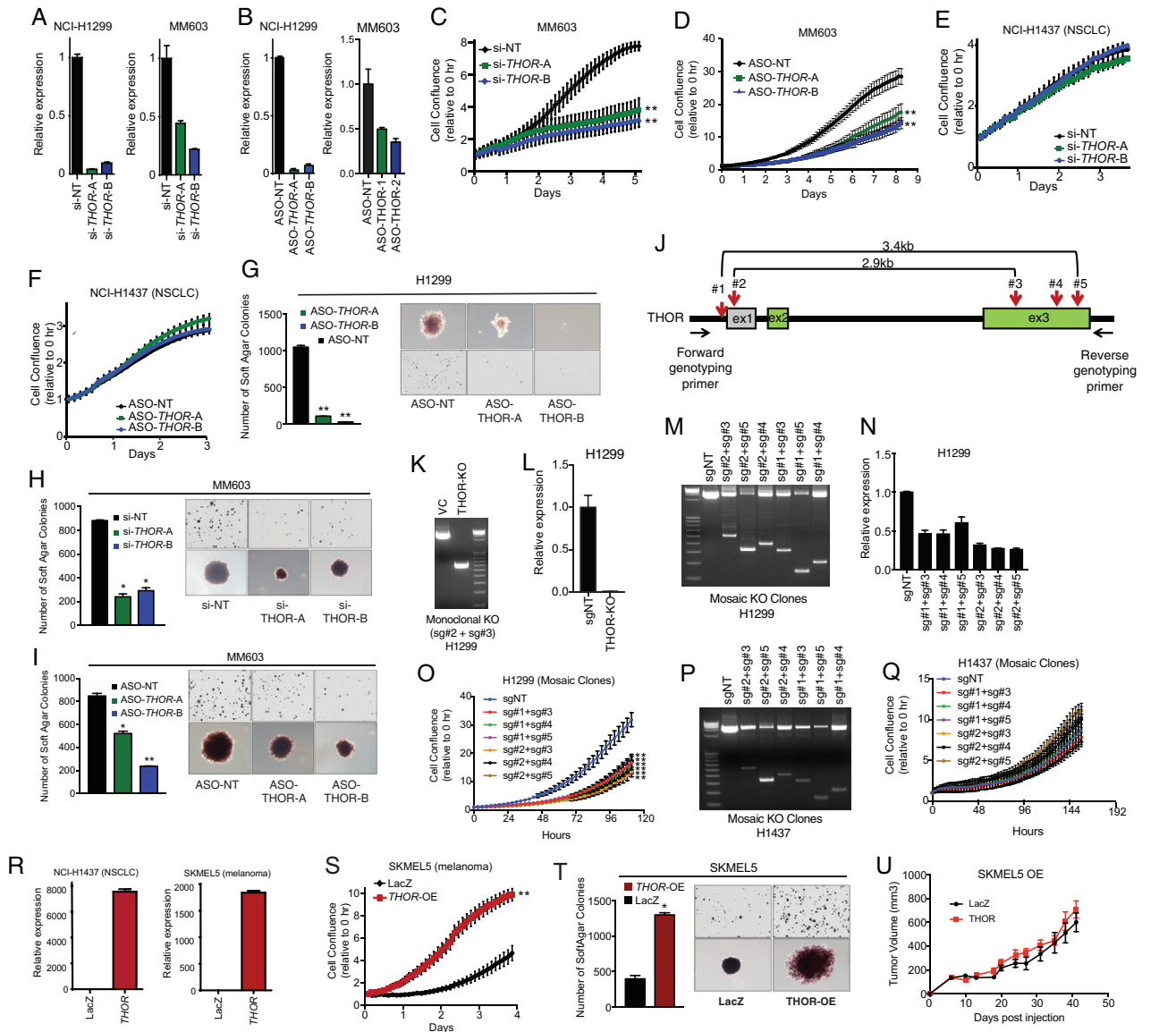(K) H&E (left) and THOR ISH (right) for the human testis, rete, and adipose.

Figure S3. *THOR* Knockdown/Knockout Efficiency and Cancer Phenotype Assays, Related to Figure 3

(A) Knockdown efficiency of two independent siRNAs against *THOR* in NCI-H1299 and MM603 cells determined by qRT-PCR. Data show mean ± SD.

(B) Knockdown efficiency of two independent ASOs against *THOR* in NCI-H1299 and MM603 cells determined by qRT-PCR. Data show mean ± SD.

(C) Cell proliferation assays for MM603 cells treated with two independent *THOR* siRNAs.

(D) Cell proliferation of MM603 cells treated with two independent ASOs .

(E) Cell proliferation assays for NCI-H1437 cells treated with two independent *THOR* siRNAs. Data show mean ± SE from one of the two independent experiments.

(F) Cell proliferation assays for SK-MEL-5 cells treated with two independent *THOR* ASOs. Data show mean ± SE from one of the two independent experiments.

(G–I) Anchorage-independent growth of (G) H1299 cells transfected with non-targeting ASO or two *THOR* ASOs, (H) MM603 cells transfected with non-targeting siRNA and siRNAs targeting *THOR*, and (I) MM603 cells transfected with non-targeting ASO and ASOs targeting *THOR*. Left, quantification of number of colonies. Right, representative image of surviving colonies and individual colony.

(J) Schematic diagram displaying the location for all guide RNAs used to generate the knockout clones. Location of forward and reverse primers also shown.

(K) DNA agarose gel confirming knockout of THOR region flanked by sg#2 and sg#3 vis PCR.

(L) qPCR validation of THOR expression in control cells compared to knockout cells.

(M) DNA agarose gel confirming knockout of regions flanked by sgRNAs in the various conditions vis PCR in H1299 cells.

(N) RNA knockout efficiency for the mosaic CRISPR knockout models determined by qPCR.

(O) Proliferation assay for the mosaic populations for the THOR knockout H1299 cells produced via various sgRNA combinations compared to non-targeting sgRNA.

(P) DNA agarose gel confirming knockout of regions flanked by sgRNAs in the various conditions vis PCR in H1437 cells.

(Q) Proliferation assay for the mosaic populations for the THOR knockout H1437 cells produced via various sgRNA combinations compared to non-targeting sgRNA.

(R) Overexpression efficiency of *THOR* in NCI-H1299 and SK-MEL-5 cells. Data show mean ± SD.

(S) Cell proliferation assay in SK-MEL-5 cells stably transfected with *THOR* overexpression or LacZ control lentivirus. Data show mean ± SE from one of the two independent experiments.

(T) Anchorage-independent growth of LacZ or *THOR* overexpressing SKMEL5 cells. Left, quantification of number of colonies. Right representative images of surviving soft agar colonies.

(U) Tumor growth for *THOR* overexpressing SKMEL5 cell line xenografts (N = 10) and control LacZ samples (N = 10). Tumor volumes at each time point by caliper measurement are shown. Asterisk (*) indicates $p \leq 0.001$ by a two-tailed Student's t test. Data show mean ± SEM from one of the two independent experiments. For all panels, asterisk (*) indicates $p \leq 0.01$, (**) indicates $p \leq 0.001$, (**) indicates $p \leq 0.0001$ by a two-tailed Student's t test.
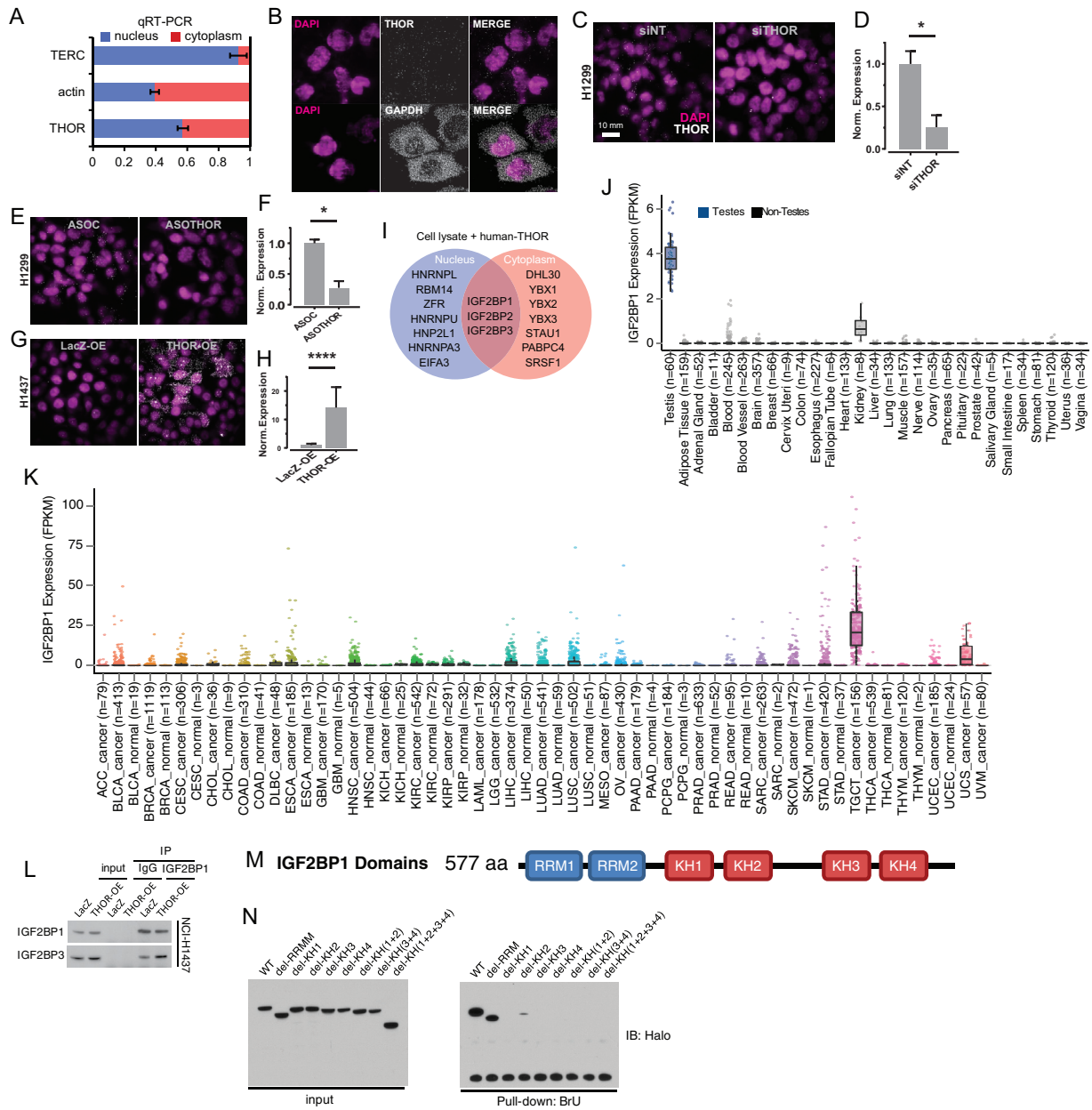
Figure S4. *THOR* Cellular Localization and Interaction with IGF2BP1, Related to Figure 4

(A) qRT-PCR for *TERC*, *ACTB*, and *THOR* following nuclear and cytoplasmic fractionation of NCI-H1299 cell lysates demonstrates both nuclear and cytoplasmic expression of *THOR*. *TERC* serves as a control for nuclear gene expression and *ATCB* serves as a control for cytoplasmic expression. Error bars represent the standard deviation (s.d.).

(B) Single molecule RNA *in situ* hybridization in NCI-H1299 cells. Staining performed for DAPI, *THOR*, and *GAPDH*.

(C–H) (C, E, G) Representative, pseudocolored images of H1299 or H1437 cells, treated with various siRNAs, ASOs or overexpression constructs and stained for DAPI (magenta) and *THOR* (gray). Scale bar, 10 μm. (D, F, H) Quantification of fold change in THOR expression of samples represented in (C), (E), and (G), respectively. Samples were normalized to siNT, ASOC and LacZ-OE respectively.

(I) Venn diagram depiction of the proteins preferentially bound to sense *THOR* (compared to antisense) from nuclear or cytoplasmic lysate from H1299 cells. Protein binding was identified via mass spectrometry following pull-down of BrU-labeled RNA.

(J) Expression of *IGF2BP1* in the GTEx normal tissue RNA-seq dataset, spanning a myriad of different normal tissue types.

(K) Expression of *IGF2BP1* across a panel of 9,714 TCGA tumor RNA-seq samples from a myriad of different tissues.

(L) western blot for IGF2BP1 and IGF2BP3 following immunoprecipitation of IGF2BP1 in the context of *THOR* or LacZ overexpression in H1437 cells.

(M) Schematic diagram of the RRM and KH domains on the IGF2BP1 protein.

(N) western blot of input (top) and following RNA-pulldown of BrU labeled *THOR* (bottom) for Halo-tagged mutant IGF2BP1 with various IGF2BP1 protein domains deleted. Error bars, s.e.m. (n = 4; > 300 cells per replicate, per sample; *p < 0.05; ****p < 0.0001).
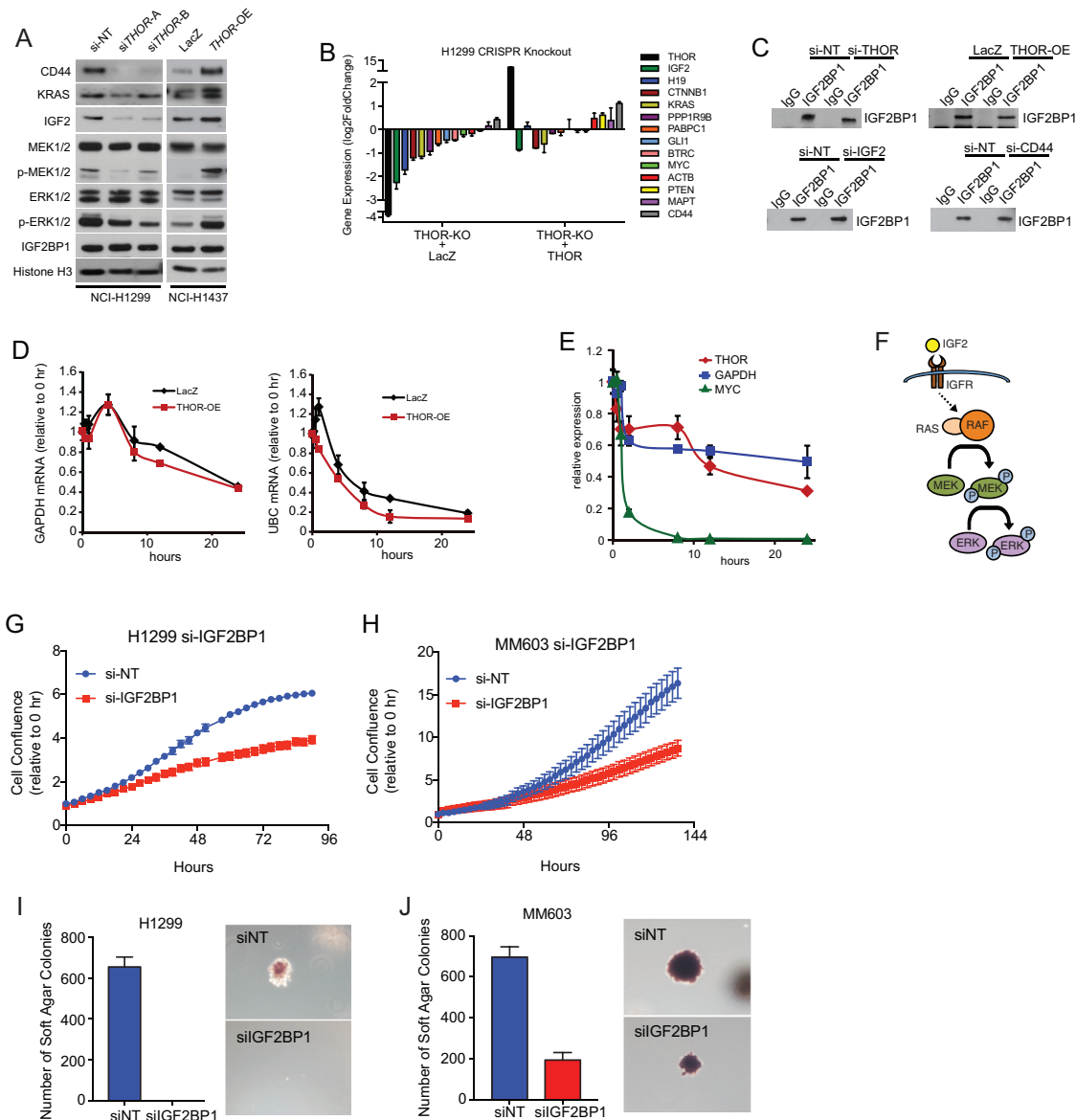
Figure S5. _THOR_ Interaction with _IGF2BP_ Proteins, Related to Figure 5

(A) The expression of IGF2BP1 targets and IGF2-MEK-ERK axis proteins and their corresponding phosphorylated forms in siTHOR treated H1299 cells (left) and _THOR_ overexpressing H1437 cells (right).

(B) Bar plot depiction of the expression levels of 13 canonical IGF2BP1 target genes by qRT-PCR in H1299 cells with CRISPR-mediated THOR knockout, and in the same cells with expression of ectopic _THOR_. Data show mean ± SD from one of the two independent experiments.

(C) western blot confirming IGF2BP1 pull-down utilized for the RIP experiments depicted in Figure 5C.

(D) qRT-PCR expression levels for _GAPDH_ (left) and _UBC_ (right) following Actinomycin D treatment in _THOR_ or _LacZ_ overexpressing H1347 cells. Data show mean ± SD from one of the two independent experiments.

(E) qRT-PCR expression levels for _THOR_ (red), _GAPDH_ (blue) and _MYC_ (green) following Actinomycin D treatment in H1299 cells. Data show mean ± SD from one of the two independent experiments.

(F) Schematic diagram of the IGF2-MEK-ERK signaling cascade.

(G and H) Cell proliferation assay for cells treated with _IGF2BP1_ siRNA in (G) H1299 cells and (H) MM603.

(I and J) Anchorage-independent growth for cells with addition of non-targeting siRNA and siRNA targeting _IGF2BP1_ in (I) H1299 and (J) MM603 cells.
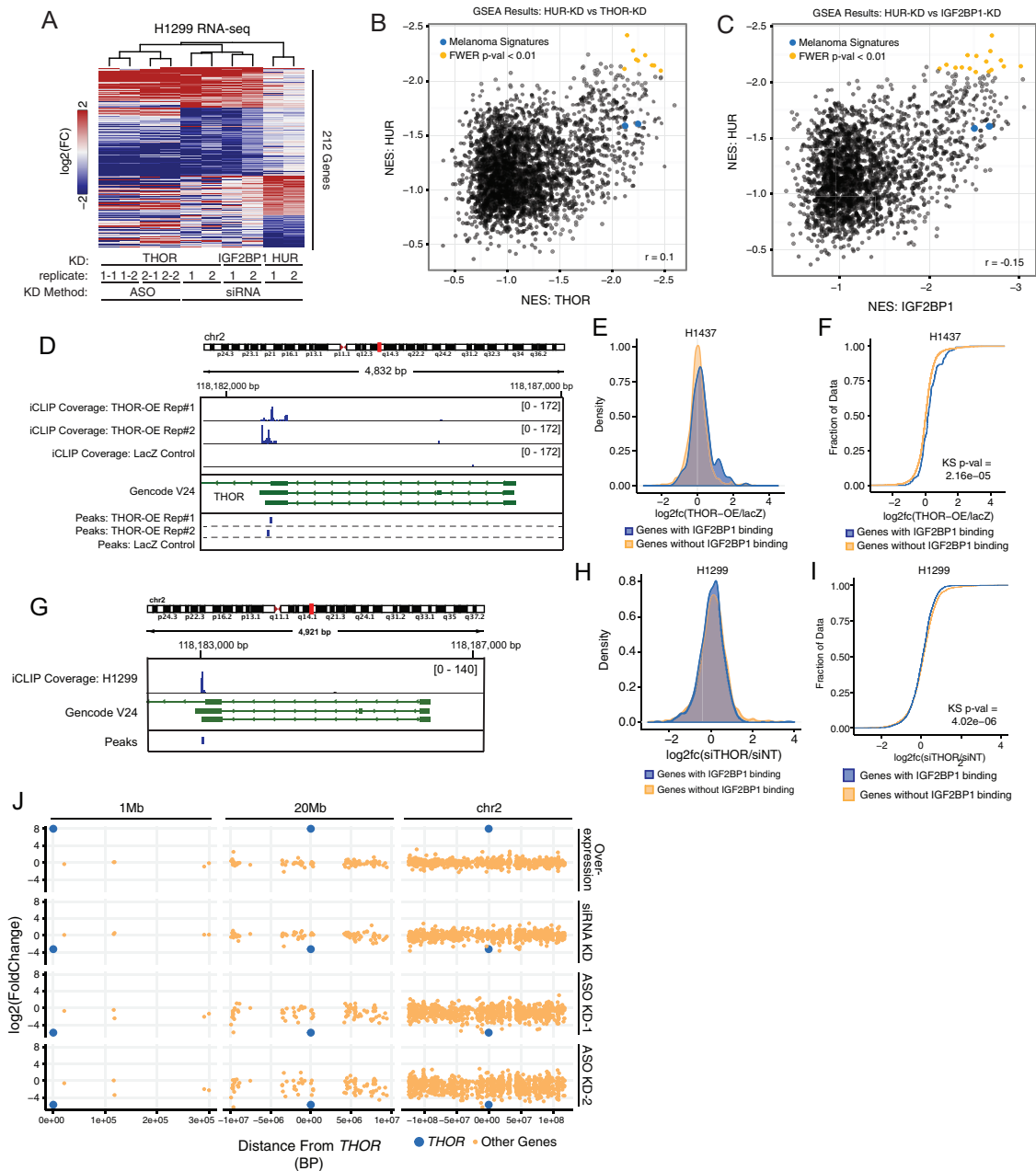
Figure S6. RNA-Seq Analysis of *THOR* Function, Related to Figure 6

(A) Heatmap depicting the expression of the genes significantly differentially expressed (DESeq FDR< 0.05) in knockdown of *THOR* and *IGF2BP1* in H1299 cells via siRNA in addition to those genes with significant differential expression in *HUR* knockdown via siRNA. Gene expression changes following ASO knockdown of *THOR* using two independent ASOs also shown. Expression depicted as the log2(fold-change) for each siRNA compared to the non-targeting siRNA control.

(B and C) Scatterplot depicting the GSEA performance for MSigDBv5.0 gene signatures with NES < 0 for (B) *HUR* and *THOR* and (C) *HUR* and *IGF2BP1*. Signatures significant upon knockdown of both genes (FWER p value < 0.01) depicted in gold.

(D) Genomic depiction of THOR. Coverage plots for IGF2BP1 replicates shown for H1437 cells overexpressing *THOR* and LacZ control (blue). GENCODEv24 gene structure of THOR also shown (green). Peaks called via Piranha for all three iCLIP samples shown (bottom).

(E and F) Gene expression depicted as log2(Fold Change) from RNA-seq data comparing the THOR-overexpression condition to LacZ overexpression. Genes identified as IGF2BP1 binding partners via iCLIP are depicted in blue, while all other genes in yellow. Expression differences shown via (E) density plot and (F) cumulative distribution function.

(G) Coverage plots for IGF2BP1 binding via iCLIP for H1299.

(H and I) Gene expression depicted as log2(Fold Change) from RNA-seq data comparing the THOR knockdown to control knockdown in H1299 cells. Genes identified as IGF2BP1 binding partners via iCLIP are depicted in blue, while all other genes in yellow. Expression differences shown via (H) density plot and (I) cumulative distribution function.

(J) Scatterplot depiction of the RNA-seq expression of genes in the genomic vicinity of THOR. Expression changes depicted as log2(Fold Change) for all conditions (*THOR* knockdown / non-targeting knockdown for the knockdown conditions and *THOR* overexpression / LacZ overexpression for the overexpression condition). Expression for genes 0.5MBP upstream, 10MBP up and downstream of THOR shown in addition to all genes on chromosome 2.
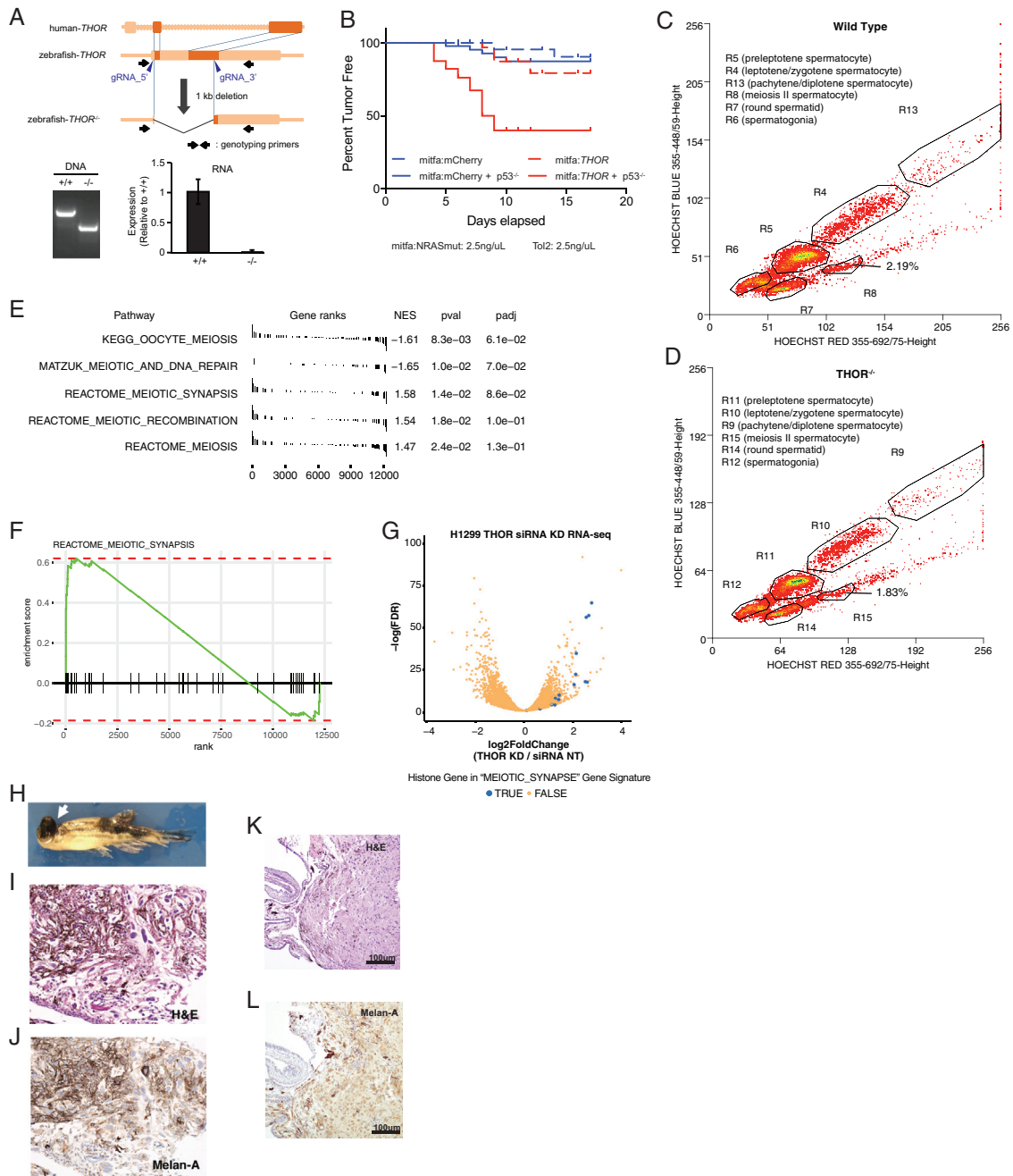
Cell

Figure S7. _THOR_ Genetic Model in Zebrafish Exhibits Melanoma Phenotype and Fertility Phenotype, Related to Figure 7

(A) Schematic representation of human and zebrafish _THOR_ transcript structure and the guide RNA (gRNA) design used to delete the conserved transcript region of zebrafish _THOR_ via CRISPR/Cas9 genome editing (top). Validation of _THOR_ knockout at both DNA (genotyping) and RNA (qRT-PCR) level (bottom). Data show mean ± SD.

(B) Kaplan-Meier curve of tumor free period for p53−/− zebrafish (solid lines) and p53 wild-type zebrafish (dotted lines) co-injected with either mitfa promoter driven NRAS 61K + mitfa promoter driven human _THOR_ (red) or mitfa promoter driven NRAS 61K + mCherry (blue).

(C and D) Cell selection is visualized in a "Hoechst Blue"/"Hoechst Red" contour plot, in which the density of the cells is displayed as contour lines that form circular contours upon high cell density. Contour plots shown for sorted zebrafish spermatocytes from (C) wild-type zebrafish and (D) _THOR_ knockout zebrafish.

(E) GSEA results shown for all 5 MSigDB signatures related to meiosis for gene expression changes following siRNA mediated _THOR_ knockdown determined by RNA-seq. Genes ranked by −log(pval)*(Fold Change).

(F) Representative GSEA plot for the REACTOME_MEIOTIC_SYNAPSE gene signature.

(G) Volcano plot for gene expression changes following _THOR_ knockdown determined via DESeq. Meiotic histone genes in the MEIOTIC_SYNAPSE gene signature shown in blue.

(legend continued on next page)

(H) Representative image of zebrafish with melanoma.

(I–L) (I and J) Immunohistochemistry for melanoma in p53 wild-type background with endogenous *THOR*. (K and L) Immunohistochemistry for melanoma in p53 knockout background zebrafish with exogenous h-THOR. (I and K) H&E staining (100x) of melanoma. (J) and (L), Immunohistochemistry staining (100x) for Melan-A of melanoma.