2-20-2018

# A *LIN28B* Tumor-Specific Transcript in Cancer

Weijie Guo
*University of Kentucky*, weijie.guo@uky.edu

Zhixiang Hu
*Fudan University, China*

Yichao Bao
*Fudan University, China*

Yuchen Li
*Fudan University, China*

Shengli Li
*Fudan University, China*

**See next page for additional authors**

**Right click to open a feedback form in a new tab to let us know how this document benefits you.**

## Repository Citation

**Authors**

Weijie Guo, Zhixiang Hu, Yichao Bao, Yuchen Li, Shengli Li, Qiupeng Zheng, Dongbin Lyu, Di Chen, Tao Yu, Yan Li, Xiaodong Zhu, Jie Ding, Yingjun Zhao, Xianghuo He, and Shenglin Huang

**A *LIN28B* Tumor-Specific Transcript in Cancer**

# A *LIN28B* Tumor-Specific Transcript in Cancer

## Graphical Abstract
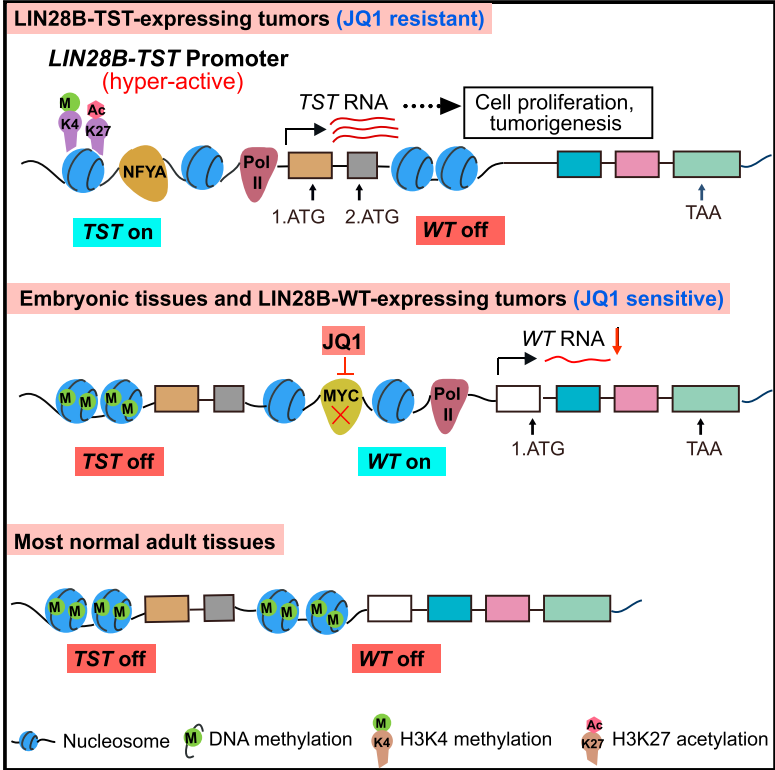


## Authors

Weijie Guo, Zhixiang Hu, Yichao Bao, ..., Yingjun Zhao, Xianghuo He, Shenglin Huang

## Correspondence

xhhe@fudan.edu.cn (X.H.), slhuang@fudan.edu.cn (S.H.)

## In Brief

Guo et al. identified a tumor-specific *LIN28B* transcript variant, *LIN28B-TST*, in hepatocellular carcinoma and many other cancer types produced by alternative transcription initiation. The *LIN28B-TST*-expressing tumors may represent a subtype of aggressive cancer. *LIN28B-TST* could serve as an ideal and promising target candidate for cancer diagnosis and therapy.

## Highlights

- RNA-seq analyses reveal a *LIN28B* tumor-specific variant in cancers

- *LIN28B-TST* initiates from an ATI site regulated by NFYA but not c-Myc

- *LIN28B-TST* expression is controlled by DNA methylation

- *LIN28B-TST* is critical for cancer cell proliferation and tumorigenesis

## Data and Software Availability

GSE77661
GSE109528
GSE109575

# A *LIN28B* Tumor-Specific Transcript in Cancer

Weijie Guo,[1,2,4,5] Zhixiang Hu,[1,2,5] Yichao Bao,[1] Yuchen Li,[1] Shengli Li,[1] Qiupeng Zheng,[1] Dongbin Lyu,[1] Di Chen,[1] Tao Yu,[3] Yan Li,[1] Xiaodong Zhu,[1] Jie Ding,[1] Yingjun Zhao,[1] Xianghuo He,[1,2,*] and Shenglin Huang[1,2,6,*]

[1]Fudan University Shanghai Cancer Center and Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, China
[2]Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China
[3]State Key Laboratory of Oncogenes and Related Genes, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200032, China
[4]Markey Cancer Center, College of Medicine, University of Kentucky, Lexington, KY 40506, USA
[5]These authors contributed equally
[6]Lead Contact
*Correspondence: xhhe@fudan.edu.cn (X.H.), slhuang@fudan.edu.cn (S.H.)
https://doi.org/10.1016/j.celrep.2018.02.002

## SUMMARY

The diversity and complexity of the cancer transcriptome may contain transcripts unique to the tumor environment. Here, we report a *LIN28B* variant, *LIN28B-TST*, which is specifically expressed in hepatocellular carcinoma (HCC) and many other cancer types. Expression of *LIN28B-TST* is associated with significantly poor prognosis in HCC patients. *LIN28B-TST* initiates from a *de novo* alternative transcription initiation site that harbors a strong promoter regulated by NFYA but not c-Myc. Demethylation of the *LIN28B-TST* promoter might be a prerequisite for its transcription and transcriptional regulation. *LIN28B-TST* encodes a protein isoform with additional N-terminal amino acids and is critical for cancer cell proliferation and tumorigenesis. Our findings reveal a mechanism of *LIN28B* activation in cancer and the potential utility of *LIN28B-TST* for clinical purposes.

## INTRODUCTION

The application of deep sequencing for RNA analysis has allowed for an unprecedented view of the human transcriptome, revealing the diversity and complexity in transcript identity and splicing (Barash et al., 2010; Djebali et al., 2012; Mercer et al., 2011). Cancer is a heterogeneous and complex disease involving genetic and epigenetic alterations that result in a more complicated transcriptome. Hence, there is tremendous potential to yield tumor-specific transcripts (TSTs) in the cancer transcriptome. The detection and characterization of TSTs such as fusion transcripts are of great importance for clinical purposes and for understanding carcinogenesis (Mertens et al., 2015; Mitelman et al., 2007). Fusion transcripts derived from DNA translocation are a typical type of TST but have only been validated in some cancers. Dysregulation of transcription splicing and epigenetic alteration may result in abundant TSTs, including specific mRNA variants and intergenic transcripts. However, the identification and functional roles of these specific transcripts in cancer remain largely unknown.

The LIN28 family, including LIN28A and LIN28B, is highly expressed during embryogenesis but silent in most adult tissues (Shyh-Chang and Daley, 2013). LIN28 can block the maturation of the tumor suppressor microRNA let-7 family and mediate diverse biological functions (Shyh-Chang and Daley, 2013; Viswanathan et al., 2008). LIN28A, in combination with NANOG, OCT4, and SOX2, can reprogram human somatic cells to pluripotent stem cells (Yu et al., 2007). LIN28A also regulates mammalian stem cell self-renewal and promotes tissue repair (Kim et al., 2014; Shyh-Chang et al., 2013). LIN28B is highly expressed in various types of human cancer (Helland et al., 2011; King et al., 2011a, 2011b; Viswanathan et al., 2009; Yang et al., 2010). Mice overexpressing *LIN28B* develop multiple tumors, including lymphoma (Beachy et al., 2012), neuroblastoma (Molenaar et al., 2012), colonic adenocarcinoma (Madison et al., 2013), Wilms tumor (Urbach et al., 2014), hepatoblastoma, and hepatocellular carcinoma (Nguyen et al., 2014). These mouse models suggest that *LIN28B* alone is sufficient to drive cancer. In addition to tumor initiation, LIN28B is necessary for maintenance of liver cancer (Nguyen et al., 2014). However, the mechanism of *LIN28B* activation in cancer remains unclear.

Here, we investigated the potential TSTs in the cancer transcriptome and found an *LIN28B* variant that is specifically expressed in hepatocellular carcinoma (HCC) and many other tumor types. We showed that *LIN28B-TST* is produced from a *de novo* alternative transcription initiation site, which harbors a strong promoter. *LIN28B-TST* encodes a long protein isoform with additional N-terminal amino acids and is critical for cancer cell proliferation and tumorigenesis.

## RESULTS

### RNA-Seq Analyses Reveal a LIN28B Tumor-Specific Variant in Multiple Cancers

We first performed transcriptome analyses of different normal and cancerous tissues, as well as several cell lines, by RNA sequencing (RNA-seq) and investigated the existence of potential TSTs (Figure S1A). We detected a substantial number of TSTs, including 430 intergenic transcripts and 1,385 variants overlapping with annotated genes (Figure 1A; Table S1). Of these TSTs, we noted an *LIN28B* variant showing a high expression level and percentage splicing index, named *LIN28B-TST*
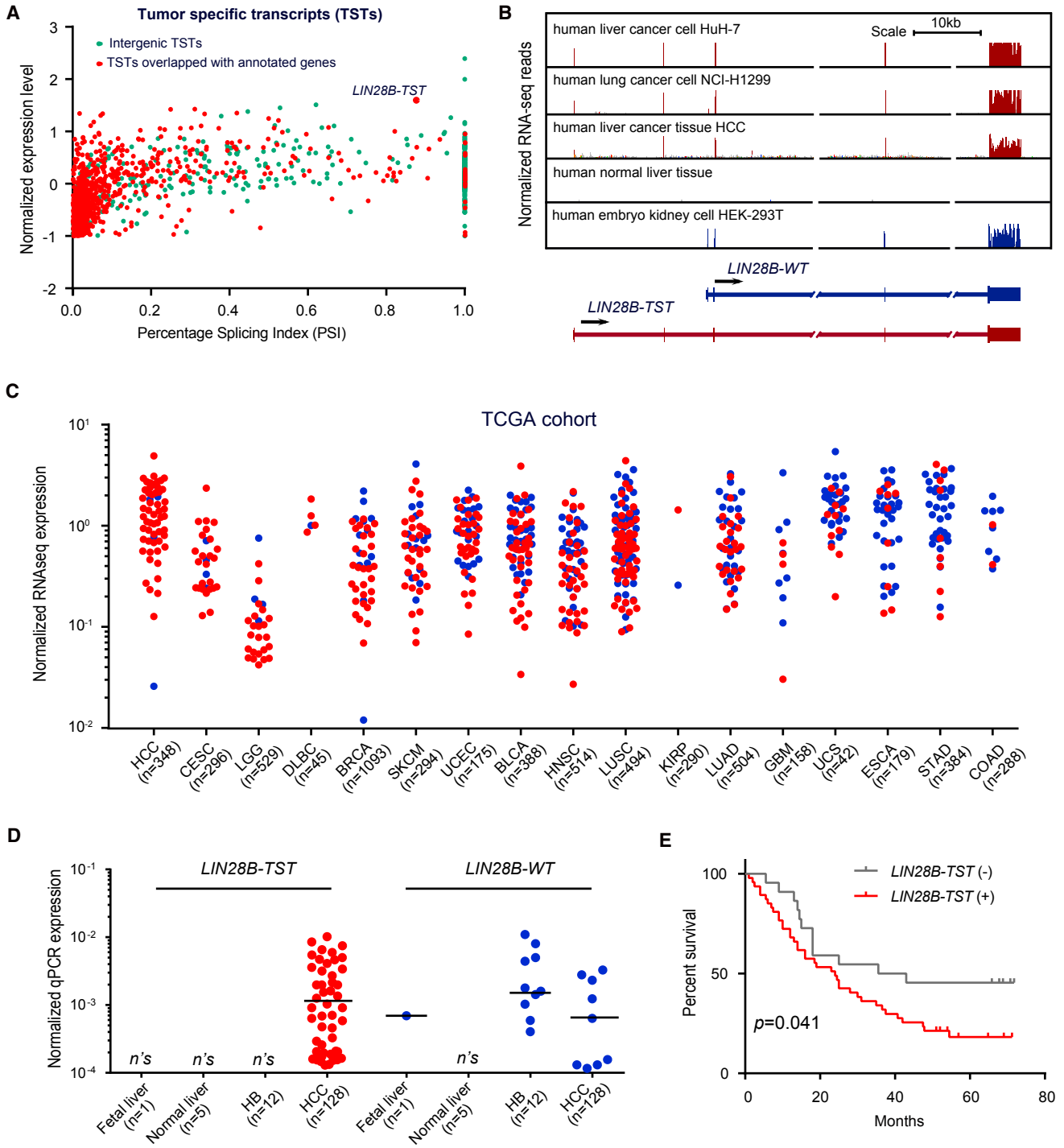
**Figure 1. RNA-Seq Analyses Revealed a *LIN28B* Tumor-Specific Transcript, *LIN28B-TST*, in Multiple Cancers**

(A) Dot plot of TST expression levels with respect to corresponding percentage splicing index (PSI) in cancer samples.

(B) Distribution of RNA-seq reads for the *LIN28B* gene locus in various RNA-seq libraries.

(C) Expression levels of *LIN28B-TST* and *LIN28B-WT* in different cancers from TCGA cohort (Table S2).

(D) Expression levels of *LIN28B-TST* and *LIN28B-WT* in fetal liver, normal liver, hepatoblastoma, and HCC. Medians are indicated within samples expressing *LIN28N-TST* or *LIN28B-WT*.

(E) Overall survival of HCC patients with (+, n = 35) or without (−, n = 22) *LIN28B-TST* expression. The p values were calculated using the log-rank test.
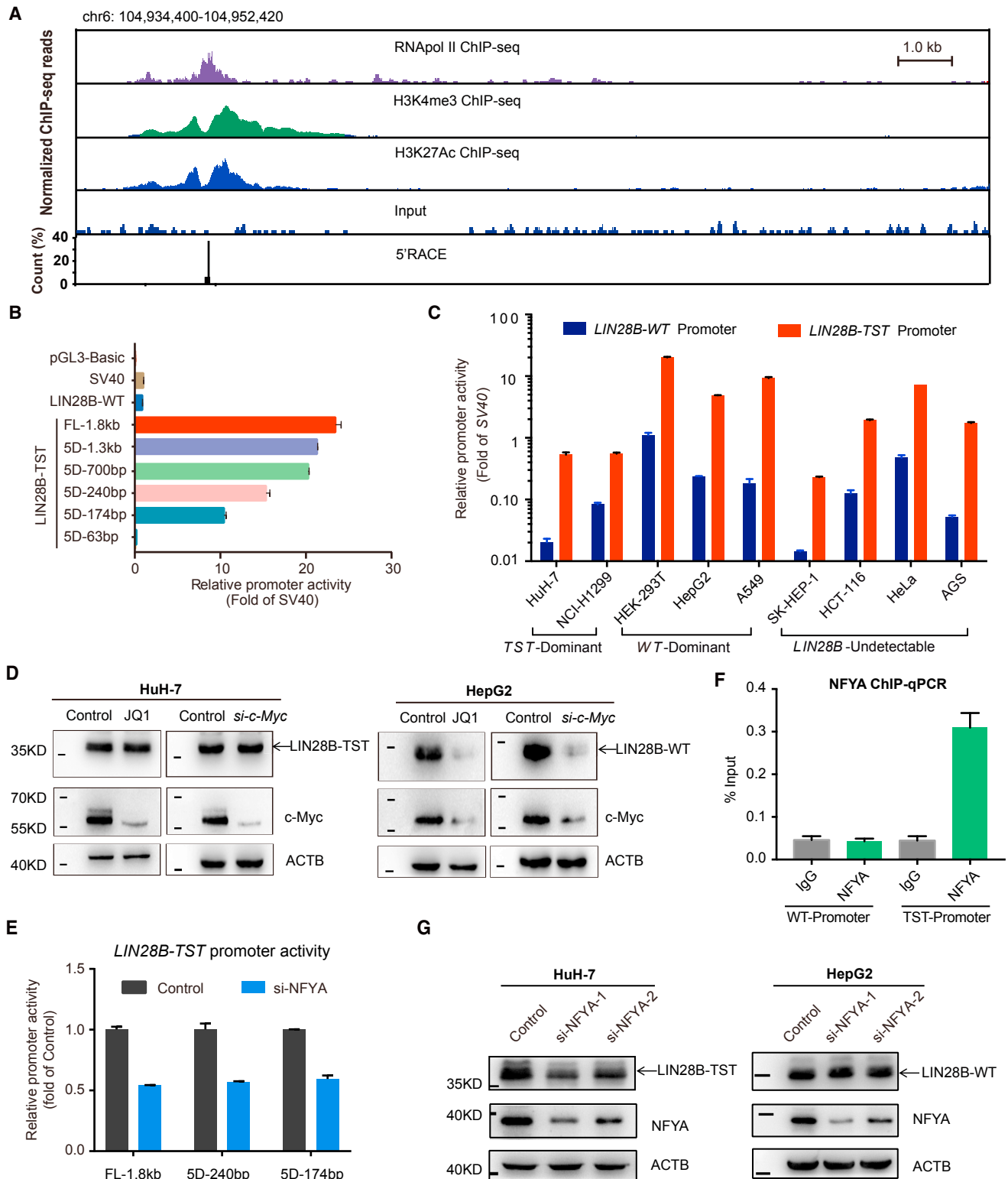
**Figure 2. *LIN28B-TST* Is Produced by Alternative Transcription Initiation with a Strong Promoter Independent of c-Myc**

(A) ChIP-seq and 5′ RACE profiles of *LIN28B-TST* at the alternative transcription initiation site in HuH-7 cells.

(B) Luciferase reporter assay of *LIN28B-TST* promoter with serial deletions from the 5′ end in HEK293T cells.

(C) Relative activity of the *LIN28B-TST* promoter in various cell lines. Data were presented as means ± SEM and representative of three independent experiments.

*(legend continued on next page)*

(Figure 1A). We found that *LIN28B-TST* was expressed in HCC tissue, HuH-7 HCC cells, and NCI-H1299 lung cancer cells, while wild-type *LIN28B* (*LIN28B-WT*) was expressed in HEK293T cells (Figure 1B). Neither variant was expressed in normal liver tissue (Figure 1B). The *LIN28B* transcript contains two new exons at the 5′ end and three exons identical to *LIN28B-WT* exons 2–4 (Figure 1B). We confirmed the presence of the *LIN28B-TST* transcript by Northern blotting and rapid amplification of cDNA ends (RACE) analysis (Figures S1C–S1F). *LIN28B-TST* differs from another alternative transcript of *LIN28B*, which was previously identified in medulloblastoma (Hovestadt et al., 2014) and expressed in H1 embryo stem cells (Figure S2A). Additionally, we found that the promoter region and new exons of *LIN28B-TST* are only conserved in primates, but not in other mammals (Figure S2B), suggesting that *LIN28B-TST* is a primate-specific variant.

To evaluate *LIN28B-TST* expression, we screened more than 7,000 tumors from 22 different cancer types in The Cancer Genome Atlas (TCGA) RNA-seq datasets. *LIN28B* was expressed in 10.2% of all cancers (Table S2). Of the *LIN28B*-expressing tumors, *LIN28B-TST* was specifically expressed in HCC (91.5%, 54/59) and many other cancer types, such as cervical and endocervical cancer (CESC; 86.2%, 25/29), brain lower grade glioma (LGG; 85.2%, 23/27), breast invasive carcinoma (BRCA; 69.8%, 30/43), skin cutaneous melanoma (SKCM; 68.1%, 32/47), and lung squamous cell carcinoma (LUSC; 50%, 59/118) (Figure 1C; Table S2). To distinguish and quantify the expression of *LIN28B-WT* and *LIN28B-TST*, we designed specific primer sets and performed qRT-PCR in a variety of cells (Figure S2C). Notably, qRT-PCR assay was more sensitive than RNA-seq analyses. We observed that *LIN28B-TST* was frequently expressed in HCC tissues (38.2%, 49/128) but not in normal or fetal liver tissues (Figure 1D). In contrast, the *LIN28B-WT* was detected in fetal liver tissue and sporadically expressed in HCC tissues (Figure 1D). Kaplan-Meier survival analysis showed that HCC patients with *LIN28B-TST* expression had poorer outcomes (Figure 1E). We also detected *LIN28B* expression in pediatric hepatoblastoma (HB) and found predominant *LIN28B-WT* expression in hepatoblastoma tissues but no *LIN28B-TST* expression (Figure 1D). Further analyses of Encyclopedia of DNA Elements (ENCODE) normal-tissue RNA-seq data showed that *LIN28B-TST* is only expressed at low levels in partial testis tissues and that *LIN28B-WT* is expressed in brain and testis tissues. Considering the immune-privileged status of the testis, *LIN28B-TST* might be regarded as a bona fide tumor-specific transcript.

## LIN28B-TST Is Produced by Alternative Transcription Initiation with a Strong Promoter

The RNA-seq profile of the *LIN28B* variant suggested an alternative transcription initiation (ATI) site approximately 20 kb upstream of *LIN28B-WT* (Figure 1B). We performed 5′ RACE and mapped the ATI site to a region on chromosome 6 at chr6:104,937,026 (hg38). Chromatin immunoprecipitation sequencing (ChIP-seq) and ChIP-qPCR revealed significant enrichment of histone H3K4me3, H3K27Ac, and RNA polymerase II (Pol II) near the ATI site (Figures 2A and S3A). These data indicated that *LIN28B-TST* originates from a newly established ATI site that is associated with characteristic chromatin alterations. To evaluate whether the region at the ATI site could function as a promoter, we cloned the putative *LIN28B-TST* promoter fragment into a luciferase reporter. The luciferase reporter assay showed that the activity of the *LIN28B-TST* promoter was over 20-fold higher than *LIN28B-WT* or the Simian vacuolating virus 40 (SV40) promoter in HEK293T cells (Figure 2B). A series of constructs with different deletions revealed a ∼200-bp core promoter proximal to the ATI site (Figure 2B). Strikingly, all cells showed strong promoter activity, including the *LIN28B-TST*-expressing cells, *LIN28B-WT*-expressing cells, and *LIN28B* null-expressing cells (Figure 2C), suggesting that transcription of *LIN28B-TST* is restricted by epigenetic mechanisms. We also investigated whether *LIN28B-TST* could be modulated by the c-Myc oncogene, which was found to directly regulate the expression of *LIN28B-WT* (Chang et al., 2009). Inhibition of c-Myc by small interference RNA (siRNA) or Bromodomain and extra terminal protein (BET) inhibitor JQ1 remarkably decreased the expression of c-Myc and LIN28B in *LIN28B-WT*-expressing HepG2 cells, while LIN28B-TST expression was not altered after silencing of c-Myc expression in *LIN28B-TST*-expressing HuH-7 cells (Figure 2E). ChIP qPCR showed that c-Myc could bind to the promoter of *LIN28B-WT* but not *LIN28B-TST* (Figure S3B), suggesting that the regulation of *LIN28B-TST* is independent of c-Myc expression. Considering that BET inhibitor JQ1 represents a promising therapeutic agent for a broad spectrum of caners and c-Myc is the key target gene, we wondered whether expression of *LIN28B-TST* is resistant to its treatment. We noted that the half maximal inhibitory concentration ($IC_{50}$) value of JQ1 in *LIN28B-TST*-expressing HuH-7 cells ($IC_{50}$ = 30.1 nM) was much higher than that in *LIN28B-WT*-expressing HepG2 cells ($IC_{50}$ = 5.08 nM) (Figures S3C and S3D). Importantly, the $IC_{50}$ value of JQ1 was significantly decreased in *LIN28B-TST*-silencing HuH-7 cells ($IC_{50}$ = 3.91 nM), almost 13% of the value in the control cells (Figure S3C). Furthermore, ectopic expression of *LIN28B-TST* led to the increase of the $IC_{50}$ value in HepG2 cells ($IC_{50}$ = 33.3 nM), compared to the control cells (Figure S3D). These findings suggested that expression of *LIN28B-TST* contributes to BET inhibitor JQ1 resistance and that inhibition of *LIN28B-TST* might be a potential strategy to sensitize LIN28B-TST-expressing cancers to JQ1.

To explore the transcription factors (TFs) that regulate the *LIN28B-TST*, we first scanned the LIN28B-TST promoter sequences, using PROMO and TFBIND, and found a panel of 28 TFs that possess putative binding sites in the *LIN28B-TST* promoter region. We then performed a siRNA knockdown screening

(D) Immunoblots of LIN28B and c-Myc expression in cells treated with 1 μM JQ1 or transfected siRNA targeting *c-Myc* for 2 days. β-actin (ACTB) was used as the loading control.

(E) NFYA enrichment near the ATI site of LIN28B-TST was assessed by ChIP qPCR in HuH-7 cells.

(F) Activities of LIN28B-TST promoter reporters with different lengths were determined in HuH-7 cells transfected with siRNAs targeting NFYA.

(G) Immunoblots of LIN28B and NFYA expression in cells transfected with siRNA targeting NFYA for 2 days. β-actin (ACTB) was used as loading control.

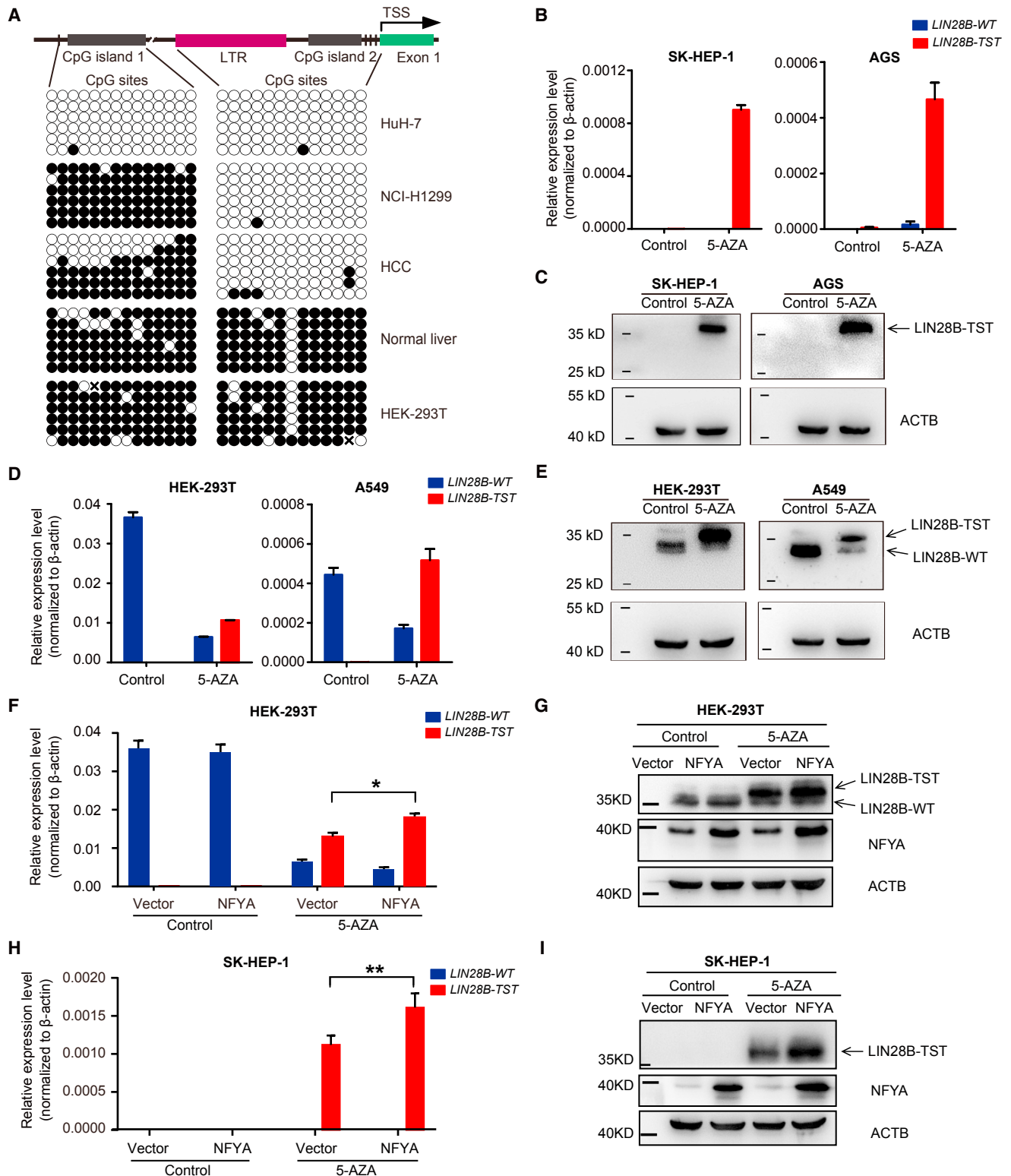**Figure 3. Demethylation of the LIN28B-TST Promoter Is Associated with Its Expression**

(A) Schematic diagram showing the position of 2 CpG islands at the promoter region of *LIN28B-TST*. Promoter methylation status was analyzed by bisulfite sequencing. Filled circles denote methylated cytosines, and open circles denote unmethylated cytosines at the indicated CpG sites.

*(legend continued on next page)*

to evaluate their effect on LIN28B-TST expression in HuH-7 cells. Of note, three TFs (NFYA, YY1, and TFAP4) were identified to regulate LIN28B-TST expression (Figure S3E), and NFYA had the most significant effect. NFYA is a pioneer transcription factor during spermatogonial stem cell development and preimplantation development (Guo et al., 2017; Lu et al., 2016). A binding site of NFYA (CCAAT) is present in the *LIN28B-TST* promoter (Figure S3F). ChIP qPCR, luciferase reporter, and western blot assays confirmed that NFYA could directly regulate the expression of *LIN28B-TST* but not *LIN28B-WT* (Figures 2F–2H).

## Demethylation of the LIN28B-TST Promoter Is Associated with Its Expression

To evaluate whether CpG methylation of the *LIN28B-TST* promoter is associated with its expression, we performed bisulfite sequencing in regions flanking the ATI site. Two CpG islands were found in the *LIN28B-TST* promoter regions (Figure 3A). Compared to *LIN28B-TST* null-expressing samples, the *LIN28B-TST*-expressing samples showed lower CpG methylation in the second CpG island (Figure 3A). To test whether demethylation at the CpG site is sufficient to induce its expression, we investigated the expression of *LIN28B-TST* in *LIN28B* null or *LIN28B-WT*-expressing cells in the presence of the DNA methyltransferase inhibitor 5-aza-deoxycytidine (5-AZA). Treatment with 5-AZA strongly activated *LIN28B-TST* mRNA and protein expression in *LIN28B* null-expressing SK-HEP-1 and AGS cells (Figures 3B and 3C). Moreover, the induction of *LIN28B-TST* significantly decreased *LIN28B-WT* mRNA and protein expression in wild-type *LIN28B*-expressing HEK293T and A549 cells (Figures 3D and 3E). We further explored whether the NFYA regulation of *LIN28B-TST* depends on its DNA methylation status. We introduced NFYA expression into *LIN28B* null or *LIN28B-WT*-expressing cells with or without 5-AZA treatment. We observed that ectopic expression of NFYA alone could not induce *LIN28B-TST* expression in these cells (Figures 3F and 3G). Interestingly, in the presence of 5-AZA, NFYA could increase the mRNA and protein levels of LIN28B-TST (Figures 3F and 3G). These results indicate that demethylation of the *LIN28B-TST* promoter might be a prerequisite for its transcription and transcriptional regulation.

## LIN28B-TST Encodes a Protein Isoform with Additional N-Terminal Amino Acids and Is Critical for Cancer Cell Proliferation and Tumorigenesis

The *LIN28B-TST* transcript contains two predicted in-frame start codons (ATGs) at the 5′ end, resulting in proteins with additional N-terminal amino acids (Figure 4A). Immunoblots of different cancer cell lines expressing *LIN28B-TST* revealed a predominant band, suggesting that *LIN28B-TST* is translated mainly from the first start codon (Figure S2D). We further mutated the two start codons individually or in combination. Immunoblots revealed that each mutated form of *LIN28B-TST* no longer pro-

duced the corresponding protein band (Figure 4B). We also prepared a rabbit antibody against the additional N-terminal amino acids. Immunoblots showed that this antibody specifically recognized the primary LIN28B-TST isoform but not the minor isoform or LIN28B-WT (Figures 4B and S2D). Using this specific LIN28B-TST antibody, we detected the expression of LIN28-TST in HCC tissues but not in normal liver tissues by immunoblotting and immunohistochemistry (Figures 4C and 4D).

Because the LIN28B-TST protein contains all of the functional domains of LIN28B-WT (Figure 4A), we speculated that LIN28B-TST would have functions similar to those of LIN28B-WT. To test this notion, we ectopically expressed LIN28B-TST or LIN28B-WT in *LIN28B* null-expressing SK-HEP-1 cells. We found that either LIN28B-TST or LIN28B-WT significantly suppressed the expression of let-7 family members (Figure 4E). Consistently, overexpression of either LIN28B-TST or LIN28B-WT increased the levels of let-7 target genes, including HMGA2, and IGF2BP1, -2, and -3. (Figure 4F). We also noted that the protein level of LIN28B-TST was higher than that of LIN28B-WT in cells infected with the same amount of lentivirus, which may be due to the different translation efficiency or protein stability. By using cycloheximide (CHX), a protein synthesis inhibitor, we observed that the stability of the LIN28B-TST protein level is much higher than that of LIN28B-WT (Figure S4A). We further used siRNAs and CRISPR/Cas9 technology to knockdown LIN28B-TST expression and observed significant inhibition of cell growth and cell-colony formation (Figures 4G and 4H; Figures S4B–S4D). Silencing of LIN28B-TST also dramatically inhibited the tumorigenic ability and reduced the tumor weight and volume of HuH-7 cells (Figures 4I and 4J). All together, these results indicated that LIN28B-TST is critical for cancer cell proliferation and tumorigenesis and may serve as a specific target for cancer treatment.

## DISCUSSION

In this study, we provide a mechanism of *LIN28B* activation through alternative transcription initiation and epigenetic alteration. Similar to the genesis of $ALK^{ATI}$, a *ALK* transcript in several cancer types (Wiesner et al., 2015), *LIN28B-TST* is also produced from a new transcription initiation site. However, these two transcripts use distinct transcriptional activation mechanisms. The $ALK^{ATI}$ expression is mainly driven by the long terminal repeat (LTR) promoter. Despite the presence of LTR within the *LIN28B-TST* promoter region, it is not considered a core component of the *LIN28B-TST* promoter. The core promoter of *LIN28B-TST* showed ubiquitously strong activity independent of cell types, and LIN28B-TST expression level correlates with its promoter methylation status, indicating that epigenetic alterations may contribute to the activation of *LIN28B-TST*.

Identifying molecules specific to tumor is essential for the deciphering of the cancer development and progression, as well as for

---

(B–E) SK-Hep1 (B and C), AGS (B and C), HEK-293T (D and E) and HepG2 (D and E). Cells were treated with or without 10 μM 5-AZA for 6 days and then harvested for qPCR and western blot assays. qPCR data were presented as means ± SEM and representative of three independent experiments.

(F–I) The LIN28B-WT-expressing HEK293T (F and G) or LIN28B null-expressing SK-HEP-1 (H and I) cells were infected with NFYA lentivirus, treated with or without 10 μM 5-AZA for 6 days, and then harvested for qPCR and western blot assays.

Data are representative of three independent experiments, and qPCR data are presented as means ± SEM. *p < 0.05; **p < 0.01.
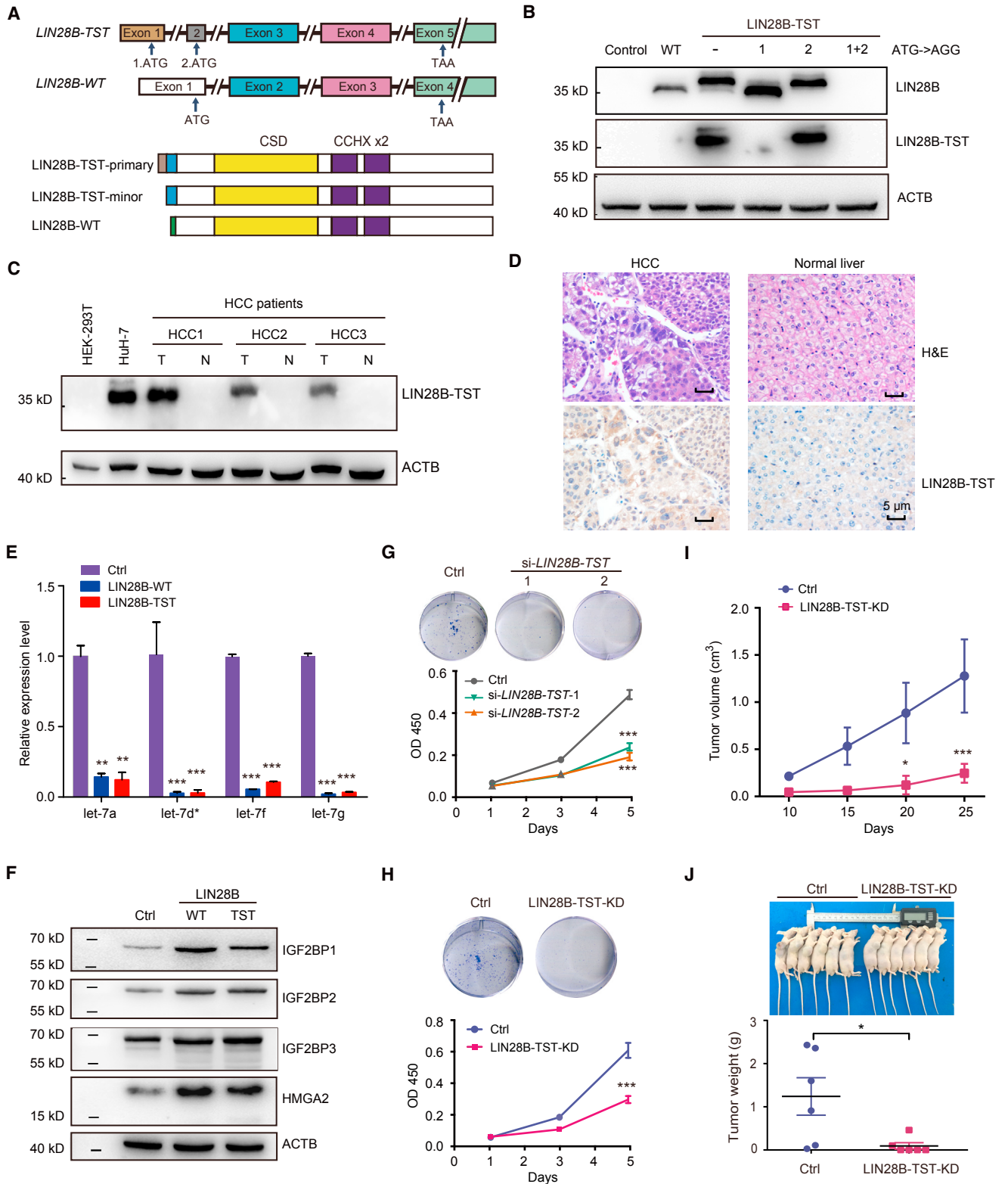
(legend on next page)

cancer diagnosis and therapy. DNA mutations constitute the major focus of tumor-specific discovery efforts to date. The diversity and complexity of cancer transcriptome hold the potential to yield TSTs in cancer. We proposed that dysregulation of transcription splicing and epigenetic alteration may result in abundant TSTs, such as specific mRNA variants and intergenic transcripts. In this study, we investigated the existence of these potential TSTs by RNA-seq and transcriptome analyses. Fusion transcripts were not included in our bioinformatics pipelines. We detected a substantial number of TSTs, suggesting widespread TSTs in cancer. Notably, we only considered the transcripts that have not been annotated previously. The number of TSTs could be underappreciated. In addition, the TST candidates could be false positive due to the limited set of samples. The analyses of large databases of transcriptome sequencing data derived from tumors and normal tissues such as TCGA and the Genotype-Tissue Expression (GTEx) project are suggested for future study.

In this study, we showed that the function of LIN28B-TST is similar to that of LIN28B-WT. However, the detailed mechanism of how the additional amino acids of LIN28B-TST may contribute to the extraordinarily high protein stability has not yet been elucidated. In addition, *LIN28B-TST* expression is independent of the c-Myc oncogene, and LIN28B-TST-expressing cells are resistant to the BRD4 inhibitor JQ1. Moreover, *LIN28B-TST* expression is associated with significantly poorer prognosis in HCC patients. The *LIN28B-TST*-expressing tumors may represent a subtype in cancer. Such tumors would be more aggressive and resistant to therapy, possibly due to the stem cell phenotype by *LIN28B* expression (Shyh-Chang and Daley, 2013). HCC is a leading cause of cancer mortality and is increasing in incidence worldwide. Chemotherapy for HCC has been quite ineffective. Our findings suggested that LIN28B-TST may serve as an ideal and promising target candidate for HCC therapy. For instance, LIN28B-TST may be targeted by epigenetic compounds, siRNA, or antisense oligonucleotide (ASO) without affecting wild-type LIN28B expression. In addition, the specific additional amino acids may be a neo-epitope candidate for tumor immunotherapy.

In conclusion, we identified a tumor-specific *LIN28B* transcript variant, *LIN28B-TST*, in HCC and many other cancers. Unlike wild-type *LIN28B*, *LIN28B-TST* initiates from a *de novo* alternative transcription initiation site harboring a strong promoter

independent of c-Myc expression. The *LIN28B-TST*-expressing tumors may represent a subtype of aggressive cancer. These findings provide new insight into the diversity and complexity of *LIN28B* transcriptional regulation, and THEY suggest a mechanism of *LIN28B* activation in cancer and the potential of *LIN28B-TST* in cancer diagnosis and therapy.

## EXPERIMENTAL PROCEDURES

Detailed materials and methods can be found in the Supplemental Experimental Procedures.

### RNA-Seq Analysis
The RNA-seq data for six normal tissues (brain, heart, liver, lung, colon, and stomach) and seven cancer tissues (hepatocellular carcinoma, colorectal cancer, gastric cancer, kidney clear cell carcinoma, bladder urothelial carcinoma, breast cancer, and prostate cancer) were previously described (Zheng et al., 2016). HuH-7, NCI-H1299, and HEK293T cells were also subjected to RNA-seq. Sequencing reads from RNA-seq data were aligned using the spliced read aligner HISAT2, v2.0.4 (Kim et al., 2015), which was supplied with the Ensembl human genome assembly (Genome Reference Consortium GRCh38) as the reference genome. Specifically, a two-step mapping strategy was used to combine splice junctions from all samples that could be used to annotate mapped reads. In the first round of mapping, HISAT2 was provided with known splice sites extracted from GENCODE annotation, v24 (Harrow et al., 2012), using hisat2_extract_splice_sites.py script. At the end of the first mapping step, all splice junctions in each sample were combined across samples to obtain a master set of combined junctions. Next, a second round of mapping was performed, with the master set of junctions generated in the first step supplied using the option "–novel-splice site-infile." The read alignments were provided as input into StringTie, v1.2.3 (Pertea et al., 2015), for transcriptome assembly. GENCODE, v24, was used as the transcript model reference annotation to guide the assembly process. Transcripts were assembled individually for each sample and then merged to generate a non-redundant set of transcripts observed in all RNA-seq samples. The Cufflinks program (Trapnell et al., 2012) was used to produce expression levels for the merged transcriptome in each sample at the gene and transcript levels in FPKM (fragments per kilobase of transcript per million mapped reads) units. The percentage splicing index of each transcript—which denotes the fraction of each transcript's abundance over its parent gene's abundance—was calculated. Only high-confidence transcripts containing at least two exons and >0.1 FPKM were considered. TSTs exclusively expressed in tumor samples were extracted by excluding all annotated transcripts (GENCODE, v24) and those expressed in normal samples.

### Cell Culture and Chemical Reagents
HEK293T, HuH-7, and SK-HEP-1 cells were cultured in DMEM; NCI-H1299 cells were cultured in RPMI-1640 medium (GIBCO); HCT-116 cells were

---

**Figure 4. *LIN28B-TST* Encodes a Protein Isoform with Additional N-Terminal Amino Acids and Is Critical for Cancer Cell Proliferation and Tumorigenesis**

(A) Schematic illustrations of *LIN28B* transcript variants and corresponding isoforms. Start codons (ATG) and stop codons (TAA) are indicated by arrows.
(B) Immunoblots of cells with expression of *LIN28B-WT* and *LIN28B-TST*. The two predicted start codons of *LIN28B-TST* were mutated from ATG to AGG, individually or in combination, as indicated.
(C) Immunoblots of HEK293T cells, HuH-7 cells, and HCC patients with the LIN28B-TST antibody. T, HCC tissue; N, matched non-tumor tissue.
(D) Immunostaining of HCC and normal liver tissues by the LIN28B-TST antibody.
(E) Expression levels of let-7 family microRNAs in SK-HEP-1 cells ectopically expressing LIN28B-WT or LIN28B-TST. Data are presented as means ± SEM and representative of three independent experiments.
(F) Immunoblots of IGFBP1-3 and HMGA2 in SK-HEP-1 cells ectopically expressing LIN28B-WT or LIN28B-TST.
(G and H) Colony formation assay and cell proliferation assay of HuH-7 cells with or without *LIN28B-TST* knockdown by siRNAs (G) or CRISPR/Cas9 (H). Two siRNAs targeting *LIN28B-TST*, si-LIN28N-TST-1 and si-LIN28B-TST-2, were used. Non-targeting siRNA was used as a negative control (Ctrl). A total of 3 small guide RNAs (sgRNAs) targeting LIN28B-TST were used, and the detailed results are presented in Figure S3. OD 450, optical density 450.
(I and J) The volume (I) and weight (J) of HuH-7 xenograft tumors with or without LIN28B-TST knockdown. Data are presented as means ± SEM (I) or median (J) (n = 6 mice per group).
*p < 0.05; **p < 0.01; ***p < 0.001. See also Figure S3.

cultured in McCoy's 5A medium (GIBCO); and AGS and A549 cells were cultured in F-12K medium (GIBCO) supplemented with 10% fetal bovine serum (GIBCO), 100 IU/mL penicillin-streptomycin (GIBCO) in a 37°C incubator with 5% $CO_2$ and a humidified atmosphere. JQ1 was purchased from Selleck Chemicals (Houston, TX, USA); CHX, and 5-AZA were purchased from Sigma-Aldrich (St. Louis, MO, USA).

### Antibodies
LIN28B-TSTs were generated by GenScript (Nanjing, China) and raised against the distinct N-terminal immunogen of LIN28B-TST (HRRQVLQKR MRSFNQVSSAP).

### RNA Extraction, Reverse Transcription, and Real-Time qPCR
Total RNA was extracted from tissues or cells using TRIzol reagent. cDNA was synthesized using the PrimeScript RT Reagent Kit (TaKaRa, Shiga, Japan). Real-time qPCR analyses were performed using SYBR Premix Ex TaqII (TaKaRa). Mature microRNAs were quantified with specific primers and probes using TaqMan MicroRNA Assays (Applied Biosystems, Foster City, CA, USA). U6 small nuclear RNA (snRNA) served as a loading control.

### Luciferase Assays
Cells were cultured in 24-well plates and co-transfected with 50 ng firefly luciferase reporter plasmid and 10 ng Renilla luciferase reporter (pRL-TK) plasmid. After 48 hr of incubation, firefly and Renilla luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega).

### DNA Methylation Analysis
Genomic DNA was extracted using the AllPrep DNA/RNA Mini Kit (QIAGEN). Two micrograms of genomic DNA were treated with sodium bisulfite using the EpiTect system (QIAGEN) following the manufacturer's protocol.

### Western Blotting
Proteins were separated on a 10% or 14% SDS-polyacrylamide gel and transferred on to a nitrocellulose membrane (Bio-Rad, Hercules, CA, USA). The membrane was blocked with 5% non-fat milk and incubated with primary antibodies.

### Transfection of Oligonucleotides
Small interfering RNAs (siRNAs) were synthesized by RiboBio (Guangzhou, China), and the sequences were listed in Table S3. Cells were transfected with the oligonucleotides using Lipofectamine RNAiMAX Reagent (Invitrogen) at a final concentration of 50 nM.

### Cell Proliferation and Colony Formation Assays
Cell proliferation was measured using the Cell Counting Kit-8 (CCK8) assay (Dojindo Laboratories, Kumamoto, Japan). For the colony formation assay, $1 \times 10^3$ cells were plated in each well of a 6-well plate and incubated at 37°C for 1–2 weeks. The cells were fixed with 4% paraformaldehyde and stained with 1% crystal violet (Sigma-Aldrich). Megascopic cell colonies were counted and analyzed.

### Xenograft in Nude Mice
HuH-7 cells with LIN28B knockdown by CRISPR/Cas9 or vector control were harvested and suspended in DMEM. Twelve mice (male BALB/c-nu/nu, 6 weeks old) were randomly divided into two groups. Each mouse was injected subcutaneously in the lower back with $2 \times 10^6$ cells in 200 μL DMEM. The mice were sacrificed after 4 weeks, and xenograft tumors were excised and weighed.

### Statistical Analysis
Data were presented as the mean ± SEM from at least three independent experiments. Unless stated otherwise, the two-tailed Student's t test or one-way ANOVA followed by Dunnett's multiple comparisons test was performed to compare differences between two groups or more than two groups, respectively. The paired t test was used to analyze mRNA expression levels in paired human samples. A p value of <0.05 was considered statistically significant.

### DATA AND SOFTWARE AVAILABILITY

The accession numbers for raw sequencing data reported in this study are GEO: GSE77661 (tissue RNA-seq), GEO: GSE109528 (cell RNA-seq), and GEO: GSE109575 (ChIP-seq).

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.02.002.

### AUTHOR CONTRIBUTIONS

S.H., X.H., and W.G. designed experiments. Z.H., W.G., Y.B., Y.L., Q.Z., D.C., T.Y., Y.L., J.D., and Y.Z. performed the experiments. S.H., W.G., S.L., Z.H., D.L., and X.Z. analyzed data. S.H., W.G., and Z.H. wrote the manuscript.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### REFERENCES

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. Nature 465, 53–59.

Beachy, S.H., Onozawa, M., Chung, Y.J., Slape, C., Bilke, S., Francis, P., Pineda, M., Walker, R.L., Meltzer, P., and Aplan, P.D. (2012). Enforced expression of Lin28b leads to impaired T-cell development, release of inflammatory cytokines, and peripheral T-cell lymphoma. Blood 120, 1048–1059.

Chang, T.C., Zeitels, L.R., Hwang, H.W., Chivukula, R.R., Wentzel, E.A., Dews, M., Jung, J., Gao, P., Dang, C.V., Beer, M.A., et al. (2009). Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation. Proc. Natl. Acad. Sci. USA 106, 3384–3389.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature 489, 101–108.

Guo, J., Grow, E.J., Yi, C., Mlcochova, H., Maher, G.J., Lindskog, C., Murphy, P.J., Wike, C.L., Carrell, D.T., Goriely, A., et al. (2017). Chromatin and single-cell RNA-seq profiling reveal dynamic signaling and metabolic transitions during human spermatogonial stem cell development. Cell Stem Cell 21, 533–546.e36.

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.

Helland, Å., Anglesio, M.S., George, J., Cowin, P.A., Johnstone, C.N., House, C.M., Sheppard, K.E., Etemadmoghadam, D., Melnyk, N., Rustgi, A.K., et al.; Australian Ovarian Cancer Study Group (2011). Deregulation of MYCN, LIN28B and LET7 in a molecular subtype of aggressive high-grade serous ovarian cancers. PLoS ONE 6, e18064.

Hovestadt, V., Jones, D.T., Picelli, S., Wang, W., Kool, M., Northcott, P.A., Sultan, M., Stachurski, K., Ryzhova, M., Warnatz, H.J., et al. (2014). Decoding the

regulatory landscape of medulloblastoma using DNA methylation sequencing. Nature 510, 537–541.

Kim, S.K., Lee, H., Han, K., Kim, S.C., Choi, Y., Park, S.W., Bak, G., Lee, Y., Choi, J.K., Kim, T.K., et al. (2014). SET7/9 methylation of the pluripotency factor LIN28A is a nucleolar localization mechanism that blocks let-7 biogenesis in human ESCs. Cell Stem Cell 15, 735–749.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360.

King, C.E., Cuatrecasas, M., Castells, A., Sepulveda, A.R., Lee, J.S., and Rustgi, A.K. (2011a). LIN28B promotes colon cancer progression and metastasis. Cancer Res. 71, 4260–4268.

King, C.E., Wang, L., Winograd, R., Madison, B.B., Mongroo, P.S., Johnstone, C.N., and Rustgi, A.K. (2011b). LIN28B fosters colon cancer migration, invasion and transformation through let-7-dependent and -independent mechanisms. Oncogene 30, 4185–4193.

Lu, F., Liu, Y., Inoue, A., Suzuki, T., Zhao, K., and Zhang, Y. (2016). Establishing chromatin regulatory landscape during mouse preimplantation development. Cell 165, 1375–1388.

Madison, B.B., Liu, Q., Zhong, X., Hahn, C.M., Lin, N., Emmett, M.J., Stanger, B.Z., Lee, J.S., and Rustgi, A.K. (2013). LIN28B promotes growth and tumorigenesis of the intestinal epithelium via Let-7. Genes Dev. 27, 2233–2245.

Mercer, T.R., Gerhardt, D.J., Dinger, M.E., Crawford, J., Trapnell, C., Jeddeloh, J.A., Mattick, J.S., and Rinn, J.L. (2011). Targeted RNA sequencing reveals the deep complexity of the human transcriptome. Nat. Biotechnol. 30, 99–104.

Mertens, F., Johansson, B., Fioretos, T., and Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. Nat. Rev. Cancer 15, 371–381.

Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nat. Rev. Cancer 7, 233–245.

Molenaar, J.J., Domingo-Fernández, R., Ebus, M.E., Lindner, S., Koster, J., Drabek, K., Mestdagh, P., van Sluis, P., Valentijn, L.J., van Nes, J., et al. (2012). LIN28B induces neuroblastoma and enhances MYCN levels via let-7 suppression. Nat. Genet. 44, 1199–1206.

Nguyen, L.H., Robinton, D.A., Seligson, M.T., Wu, L., Li, L., Rakheja, D., Comerford, S.A., Ramezani, S., Sun, X., Parikh, M.S., et al. (2014). Lin28b is sufficient to drive liver cancer and necessary for its maintenance in murine models. Cancer Cell 26, 248–261.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290–295.

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

Shyh-Chang, N., Zhu, H., Yvanka de Soysa, T., Shinoda, G., Seligson, M.T., Tsanov, K.M., Nguyen, L., Asara, J.M., Cantley, L.C., and Daley, G.Q. (2013). Lin28 enhances tissue repair by reprogramming cellular metabolism. Cell 155, 778–792.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–578.

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

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

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

Wiesner, T., Lee, W., Obenauf, A.C., Ran, L., Murali, R., Zhang, Q.F., Wong, E.W., Hu, W., Scott, S.N., Shah, R.H., et al. (2015). Alternative transcription initiation leads to expression of a novel ALK isoform in cancer. Nature 526, 453–457.

Yang, X., Lin, X., Zhong, X., Kaur, S., Li, N., Liang, S., Lassus, H., Wang, L., Katsaros, D., Montone, K., et al. (2010). Double-negative feedback loop between reprogramming factor LIN28 and microRNA let-7 regulates aldehyde dehydrogenase 1-positive cancer stem cells. Cancer Res. 70, 9463–9472.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917–1920.

Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., Chen, B., Luo, Y., Lyu, D., Li, Y., Shi, G., et al. (2016). Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. Nat. Commun. 7, 11215.