nature structural & molecular biology

NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing

Li Jiang[1,7], Changwei Shao[1,2,7], Qi-Jia Wu[1,3,7], Geng Chen[1], Jie Zhou[1], Bo Yang[1], Hairi Li[2], Lan-Tao Gou[2], Yi Zhang[1,3], Yangming Wang[4], Gene W Yeo[2], Yu Zhou[1,5] & Xiang-Dong Fu[1,2,6]

MicroRNA (miRNA) biogenesis is known to be modulated by a variety of RNA-binding proteins (RBPs), but in most cases, individual RBPs appear to influence the processing of a small subset of target miRNAs. Here, we report that the RNA-binding NONO–PSF heterodimer binds a large number of expressed pri-miRNAs in HeLa cells to globally enhance pri-miRNA processing by the Drosha–DGCR8 Microprocessor. NONO and PSF are key components of paraspeckles organized by the long noncoding RNA (lncRNA) NEAT1. We further demonstrate that NEAT1 also has a profound effect on global pri-miRNA processing. Mechanistic dissection reveals that NEAT1 broadly interacts with the NONO–PSF heterodimer as well as many other RBPs and that multiple RNA segments in NEAT1, including a 'pseudo pri-miRNA' near its 3′ end, help attract the Microprocessor. These findings suggest a 'bird nest' model in which an lncRNA orchestrates efficient processing of potentially an entire class of small noncoding RNAs in the nucleus.

miRNAs are a class of 21–22 nt small noncoding RNAs that are extensively involved in post-transcriptional regulation of gene expression in diverse organisms from plant to animal[1–3]. Whereas a small fraction of miRNAs are encoded by their own genes, ~80% of annotated miRNAs are derived from various large coding and noncoding transcripts[4]. These initial transcripts, known as pri-miRNAs, are pro cessed to pre-miRNAs by the Microprocessor, consisting of DROSHA and DGCR8 in the nucleus; after nuclear export by Exportin 5, pre-miRNAs are further processed into mature miRNAs by DICER before entering the RNA-induced silencing complex, or RISC[5].

Although the general miRNA biogenesis pathway has been elucidated and the core machineries for each processing step have been identified and well characterized, it is also known that each step of miRNA biogenesis is subjected to modulation, resulting in homeostatic expression of miRNAs in a highly cell-type- and tissue-specific manner. A recent study revealed a key histone variant involved in the global regulation of pri-miRNA expression at the level of transcription[6]. Pri-miRNA processing has been suggested to take place cotranscriptionally, but the existing evidence is largely based on the characterization of one or a few pri-miRNAs; however, these characterizations were not conducted genome-wide[7–10]. Therefore, although cotranscriptional pri-miRNA processing has been a popular and, to a large extent, widely accepted concept, we actually do not know to what degree a given pri-miRNA is processed cotranscriptionally or post-transcriptionally, which is an important question for understanding miRNA biogenesis and regulation in specific cells, tissues, and organs under normal physiological conditions and during disease processes.

During and after transcription, a large number of RBPs and RNA helicases, as well as post-translational modifications of these regulators, have been documented to modulate miRNA biogenesis in individual processing steps[5]. To date, however, all characterized RBPs appear to modulate a single miRNA or a small subset of miRNAs by interacting with specific *cis*-acting elements and/or secondary structures in pri- or pre-miRNAs. Such specific modulation probably contributes to the differential expression of miRNAs, even among those expressed from the same pri-miRNA transcripts.

Paraspeckles were discovered in 2002 through the identification of specific RBPs that are localized adjacent to nuclear speckles, where most pre-mRNA-processing factors are concentrated[11,12]. Both speckles and paraspeckles are permanent nuclear subdomains in most cell types, but their functions have been a continuous subject for debate and investigation[13–18]. Notably, each of these nuclear subdomains is associated with an abundant lncRNA known as MALAT1 in speckles and NEAT1 in paraspeckles[19,20], with NEAT1, but not MALAT1, being essential for maintaining the structural integrity of the corresponding nuclear subdomain[17,21,22]. Thus far, the only known or postulated function for paraspeckles is in the retention of certain

[1]State Key Laboratory of Virology and Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan, China. [2]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California, USA. [3]Present addresses: Seqhealth Technology Co., Ltd, Wuhan, China (Q.-J.W.); ABLife Inc., Wuhan, China (Y.Z.). [4]Institute of Molecular Medicine, Peking University, Beijing, China. [5]Institute for Advanced Studies, Wuhan University, Wuhan, China. [6]Institute of Genomic Medicine, University of California, San Diego, La Jolla, California, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to Y.Z. (yu.zhou@whu.edu.cn) or X.-D.F. (xdfu@ucsd.edu).

Received 17 December 2016; accepted 1 August 2017; published online 28 August 2017; doi:10.1038/nsmb.3455

hairpin-containing RNAs, particularly those derived from expressed Alu repeats[23,24], and another specific lncRNA[23], and in the sequestration of various RBPs[25,26]. However, both MALAT1 and NEAT1 have been shown to interact with some actively transcribed genes in the nucleus[15,27], thus raising the question of whether these lncRNAs and/or their associated nuclear subdomains are more actively involved in regulated gene expression, rather than simply serving as storage sites for various RNAs and proteins.

In the present study, we initially examined the differential expression of miRNAs processed from the same pri-miRNAs, which led to the elucidation of roles for NEAT1 and key paraspeckle components in the global regulation of pri-miRNA processing. Interestingly, NEAT1 harbors an apparently pseudo miRNA, which is poorly processed into mature miRNA. We found that this pseudo miRNA functions to attract the Microprocessor, whereas other RNA sequences and/or secondary structures in NEAT1 provide a general binding platform for various RBPs, some of which are engaged in extensive interactions with expressed pri-miRNAs. Our findings suggest a bird nest model for an lncRNA-organized machinery to globally enhance pri-miRNA processing, which also reveals critical insights into the formation and function of paraspeckles in the nucleus.

RESULTS

Paraspeckle components involved in pri-miRNA processing

We initially wanted to understand how different miRNAs encoded in the same primary transcripts were differentially processed in the cell. For instance, the primary miR-17-92a transcript gave rise to six mature miRNAs with dramatic differences in abundance in HeLa cells (**Supplementary Fig. 1a**; primers used for quantitative analysis are in **Supplementary Table 1**). Neither the knockdown of DICER nor the combined knockdown of AGO proteins 1–4 altered the relative abundance of individual miRNAs from the *pri-miR-17-92a* locus (**Supplementary Fig. 1b–e**; antibodies used for western blotting are in **Supplementary Table 2**), thus implying differential miRNA processing at the pri-miRNA level, which is known to be modulated by various RBPs[5]. We therefore prepared individual biotinylated pri-miRNAs from the *miR-17-92a* locus to compare their relative efficiencies in pulling down specific proteins from HeLa nuclear extracts. Pri-miR-19a and pri-miR-19b appeared to be more efficient in pulling down several proteins, which we identified by means of MS to correspond to two classes of RBPs (**Fig. 1a**; peptides identified by MS are listed in **Supplementary Table 3**). We confirmed the binding of these proteins on multiple pri-miRNAs by means of western blotting (data not shown) and direct *in vivo* crosslinking. One of the classes of RBPs contained NONO (also known as P54NRB), PSF (SFPQ), and PSPC1 (PSP1), all of which are key RBP constituents of paraspeckles[28]. The other class consisted of ILF3 (NF90) and ILF2 (NF45), which were previously implicated in nuclear export of a viral RNA[29]. We also identified heterogeneous ribonucleoprotein hnRNP A2/B1 and hnRNPA1, the latter of which has been previously shown to enhance pri-miR-18a processing[30]. Because pri-miRNAs are hairpin-containing RNAs, we chose to focus on the RBPs associated with paraspeckles whose sole function known to date is to retain or sequestrate various Alu-derived hairpin RNAs, lncRNAs, and RBPs in the nucleus[23–26,31].

Using RT-qPCR, we first determined whether the individual paraspeckle-associated RBPs we identified might affect the production of mature miRNAs. Using two independent short interfering RNAs (siRNAs) against each RBP, we found that knockdown of NONO and PSF, but not PSPC1, reduced the expression of all miRNAs from the *miR-17-92a* locus and showed corresponding increases in their

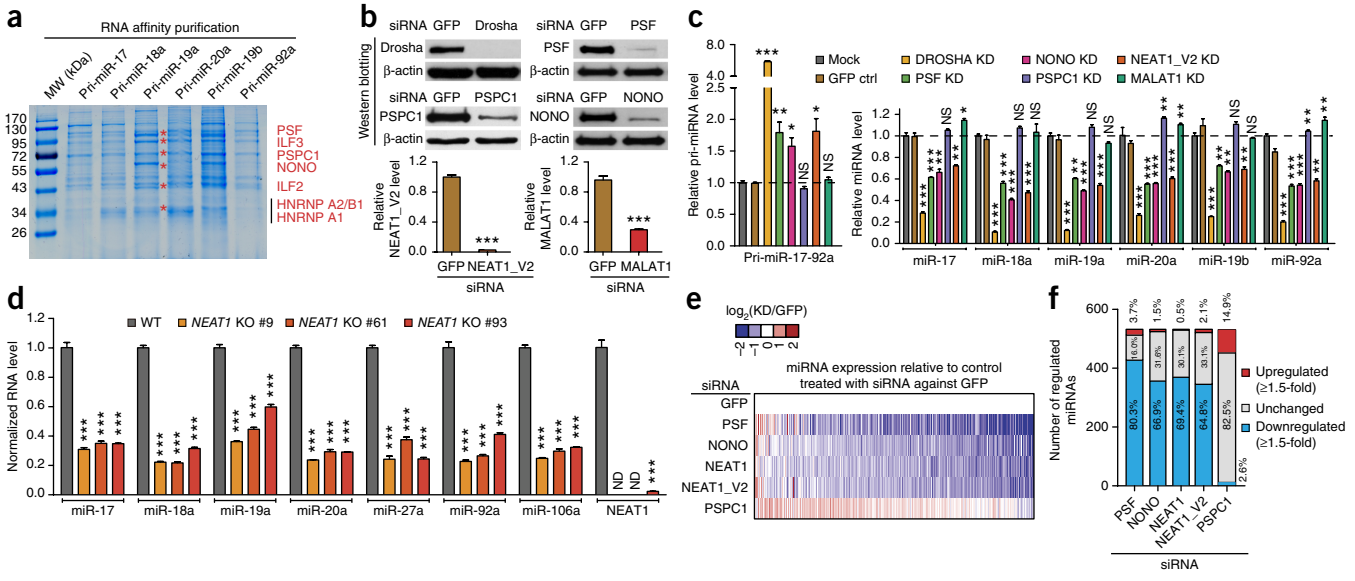

Figure 1 Involvement of paraspeckle-associated proteins and lncRNA in pri-miRNA processing. (**a**) Coomassie brilliant blue staining of proteins captured by individual pri-miRNAs from the *pri-miR-17-92a* locus from HeLa nuclear extracts. Specific proteins identified by MS, denoted by asterisks, are indicated on the right. (**b**) Knockdown of three paraspeckle-associated proteins and transcripts, NEAT1_V2 or MALAT1, quantified by western blotting and RT-qPCR, respectively. GFP indicates control treated with siRNA against GFP. Uncropped images of western blots are shown in **Supplementary Data Set 1**. (**c**) The expression of pri-miR-17-92a (left) and individual mature miRNAs from the *pri-miR-17-92a* locus (right) in response to knockdown (KD) of paraspeckle-associated factors and NEAT1, determined by RT-qPCR. "Mock" refers to cells treated with only transfection reagents. (**d**) RT-qPCR analysis of miRNAs and NEAT1 in *NEAT1*-knockout (KO) cells. Data in **b**,**c**, and **d** represent mean ± s.e.m. ($n = 3$ technical replicates). $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant, determined by two-tailed Student's t test. ND, not detectable. WT, wild-type cells. (**e**) miRNA profiling in response to specific knockdowns as in **b**, relative to a control treated with siRNA against GFP. Color key on top indicates changes in $\log_2$ scale. (**f**) Summary of the number of upregulated (≥1.5-fold), unchanged or downregulated (≥1.5-fold) miRNAs based on small RNA-seq in response to specific knockdowns as in **e**. Source data for graphs are available online.

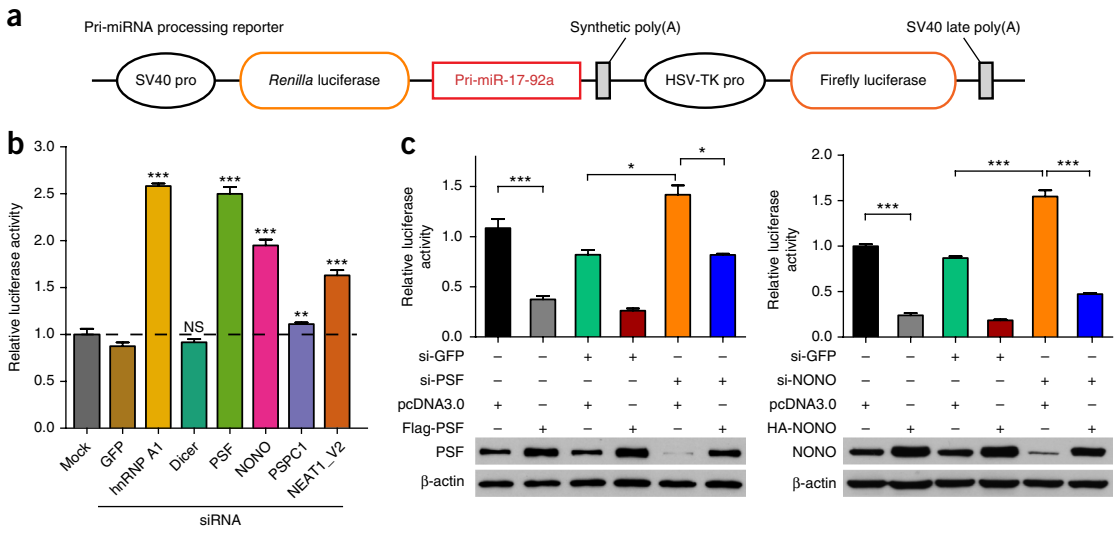

Figure 2 Function of NONO–PSF and NEAT1 analyzed with the pri-miRNA-processing reporter. (**a**) Overview of the pri-miRNA-processing reporter with pri-miR-17-92a cloned into the 3′ UTR of *Renilla* luciferase. (**b**) Relative luciferase activities of the pri-miRNA-processing reporter in response to knockdown of individual RBPs as indicated or NEAT1_V2. (**c**) Relative luciferase activities of the pri-miRNA-processing reporter in response to overexpression of siRNA-resistant PSF (left) or NONO (right). Bar graphs in **b** and **c** represent mean ± s.e.m. ($n = 3$ independent cell cultures). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. NS, not significant, determined by two-tailed Student's t test. Source data for graphs are available online. Uncropped images of western blots in **c** are shown in **Supplementary Data Set 1**.

pri-miRNA in HeLa cells (**Fig. 1b**,**c** and **Supplementary Fig. 2a**–**c**). We further confirmed these results using individual miRNA-sensor reporters (**Supplementary Fig. 2d**,**e**). We noted that the effects of NONO and PSF knockdowns were relatively weak compared to those of DROSHA knockdown, thus implying that NONO and PSF show a degree of positive influence on pri-miRNA processing, rather than being essential for the process. However, we could not rule out the possibility that the residual proteins provided part of the essential function. This is particularly pertinent to PSPC1, as knockdown of PSPC1 caused a minor increase in the expression of multiple miRNAs (**Fig. 1c**, right). We therefore attempted to use CRISPR–Cas to generate knockout cell lines for each of these RBPs. Knockout of *NONO* or *PSF* caused cell lethality, consistent with their involvement in many critical cellular functions, including transcription and pre-mRNA splicing[32–34]. In contrast, PSPC1 appeared to be dispensable for cell viability. Using two independent *PSPC1*-null cell lines (the absence of detectable PSPC1 protein confirmed by western blotting), we found that ablation of *PSPC1* significantly increased the expression of multiple miRNAs that we examined (**Supplementary Fig. 2f**,**g**), an effect also shown from the partial knockdown of PSPC1 by siRNA (**Fig. 1c**). Because NONO and PSF, but not PSPC1, are required for the structural integrity of paraspeckles[22,28] (**Supplementary Fig. 3a**), we chose to first focus on understanding the mechanism for these two paraspeckle components stimulating pri-miRNA processing.

Global effect of NONO–PSF and NEAT1 in pri-miRNA processing

Given the correlation between paraspeckle disassembly and compromised pri-miRNA processing in NONO- and PSF-knockdown cells, we were curious about the role of NEAT1, which is required for the organization and maintenance of paraspeckles[21,22,24,35]. We observed a similar effect on pri-miRNA processing upon knockdown of NEAT1, and by contrast, knockdown of MALAT1, associated with nuclear speckles[36,37], showed no effect (**Fig. 1b**,**c**). We obtained effects similar to those of NEAT1 knockdown by using a 'stealth' siRNA, in which the sense strand is modified so that only the antisense strand can enter the RISC to minimize potential off-target effects[26], and by generating *NEAT1*-null cells with CRISPR–Cas (**Fig. 1d** and **Supplementary Fig. 3b**). None of the knockdowns had a measurable influence on the expression of the Microprocessor DROSHA–DGCR8 or multiple other paraspeckle-associated RBPs we examined (**Supplementary Fig. 3c**–**f**). These data revealed key roles of specific paraspeckle components, as well as the organizing lncRNA, in enhancing miRNA biogenesis at the pri-miRNA level.

The data presented above were based on analysis of miRNAs from the *mir-17-92a* locus. To explore the potential functional impact genome-wide, we performed small-RNA sequencing (small-RNA-seq) in response to knockdown of NONO, PSF, or PSPC1 in HeLa cells. Because human *NEAT1* expresses two isoforms, V2 (23-kb full-length NEAT1, 20 kb in mice) and V1 (3.7-kb 5′ portion of V2, 3.2 kb in mice), the latter of which results from an early polyadenylation event[18,35,38], we separately knocked down NEAT1_V1 and NEAT1_V2 (note that V1 knockdown would also diminish V2, and thus, we label it as NEAT1). We performed small-RNA-seq under each treatment condition in duplicate and included a spike-in RNA during library construction for quantitative analysis (Online Methods, sequencing statistics shown in **Supplementary Table 4**). The amount of spike-in RNA, as well as fragments of other noncoding RNAs, such as tRNAs, snoRNAs, and rRNAs, showed a linear relationship between duplicated experiments, despite different sequencing depths (**Supplementary Fig. 4a**,**b**); upon normalization against both external and internal reference RNAs[39], all duplicated libraries showed high reproducibility (**Supplementary Fig. 4c**,**d**). We thus combined uniquely mapped reads from duplicated libraries, obtaining ~20 million total uniquely mapped reads under each treatment condition. We plotted the miRNA levels from each knockdown against those from control cells treated with siRNA against GFP. Strikingly, 64–80% of a total of 532 expressed miRNAs with a read number >30 in control siRNA-treated HeLa cells were down regulated upon knockdown of NONO, PSF, and NEAT1 (both V1 and V2), and again, PSPC1 knockdown showed the opposite effect on many miRNAs (**Fig. 1e**,**f**). We validated the sequencing results

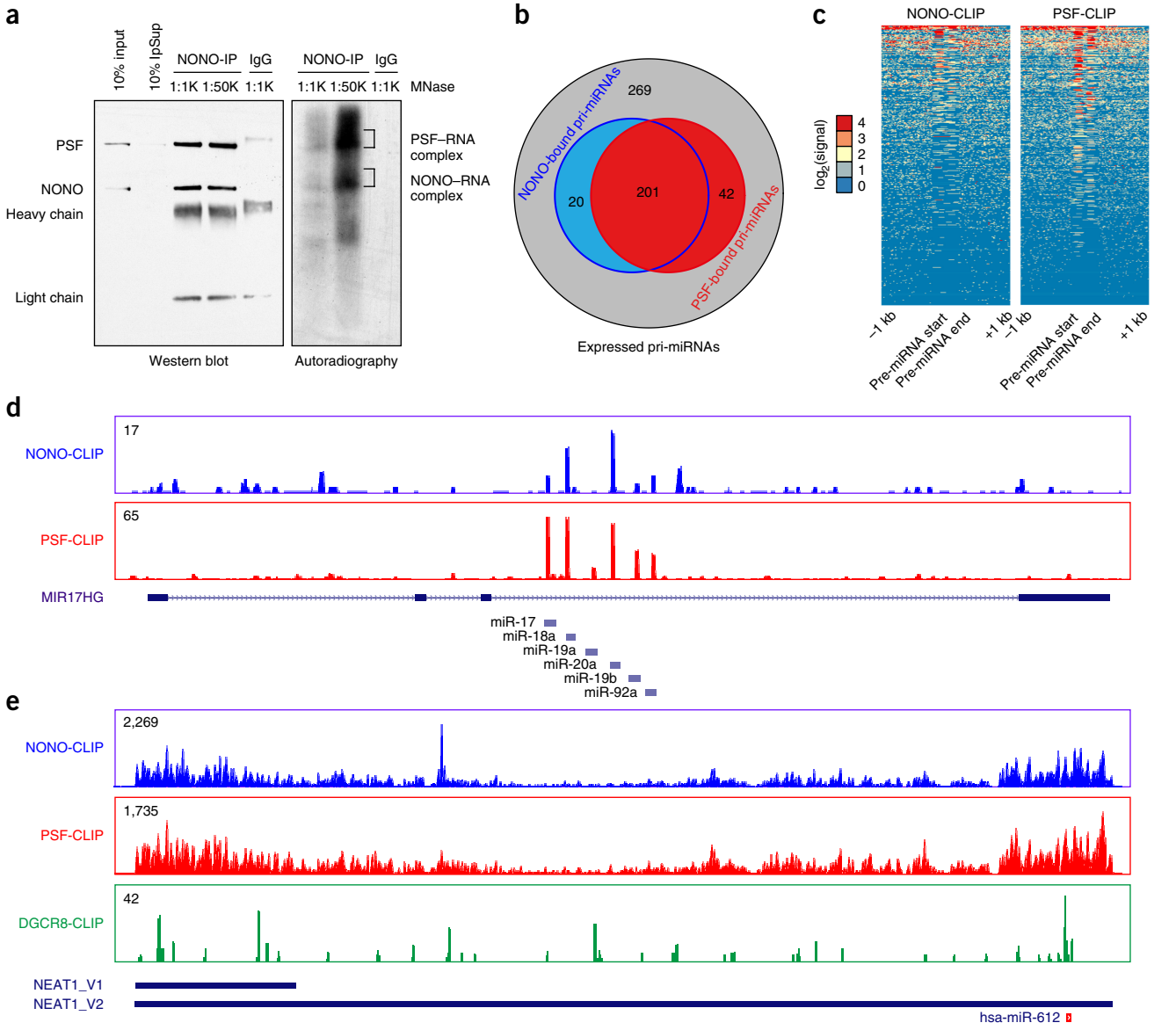

Figure 3 Genome-wide analysis of NONO–PSF–RNA interactions. (**a**) Immunoprecipitated NONO–PSF crosslinked to RNA. The complex was treated with two different concentrations of MNase (micrococcal nuclease; 1:1,000 or 1:50,000 dilution); RNA in the complex was ^{32}P labeled with T4 polynucleotide kinase. Proteins and RNA were visualized by western blotting (left) and autoradiography (right). Indicated bands (right) were individually isolated for CLIP-seq library construction. Uncropped images are shown in **Supplementary Data Set 1**. (**b**) Venn diagram showing overlapped pri-miRNAs bound by NONO and PSF in HeLa cells. (**c**) Footprint of NONO and PSF on pri-miRNAs. (**d**) Representative NONO and PSF binding tracks on the pri-miR-17-92a transcript. (**e**) Binding profiles of NONO and PSF on NEAT1 in comparison with published DGCR8 CLIP-seq signals[44]. *y* axes in **d** and **e** show CLIP-seq read density. The region encoding miR-612 is indicated at bottom.

by means of RT-qPCR on a large panel of miRNAs, (**Supplementary Fig. 4e**,**f**). These findings revealed a global role of specific paraspeckle-associated RBPs and NEAT1 in pri-miRNA processing.

To further demonstrate compromised pri-miRNA processing, we constructed a pri-miRNA-processing reporter by inserting the pri-miR-17-92a sequence in the 3′ UTR of the *Renilla* luciferase reporter (**Fig. 2a**); compromised pri-miRNA processing would lead to increased luciferase activity. Knockdown of NONO, PSF, or NEAT1 all caused elevated luciferase activities, similar to knockdown of HNRNPA1, as previously shown[30], whereas knockdown of either DICER or PSPC1 showed no effect (**Fig. 2b**). These observations imply a more direct role of NONO–PSF, and NEAT1, but not PSPC1, in pri-miRNA processing. We next performed overexpression–rescue experiments and demonstrated that overexpression of NONO and PSF stimulated pri-miRNA processing, and their siRNA-resistant cDNAs rescued the defects in specific siRNA-treated cells (**Fig. 2c**). These data provide further evidence for the involvement of NONO–PSF in enhancing pri-miRNA processing by the Microprocessor.

Prevalent binding of NONO–PSF on expressed pri-miRNAs in HeLa cells

To investigate how NONO–PSF might facilitate global pri-miRNA processing, we performed UV crosslinking immunoprecipitation coupled with deep sequencing (CLIP-seq) to identify their direct RNA targets. Both anti-NONO and anti-PSF antibodies efficiently brought down the NONO–PSF heterodimer, as reported previously[40,41], each

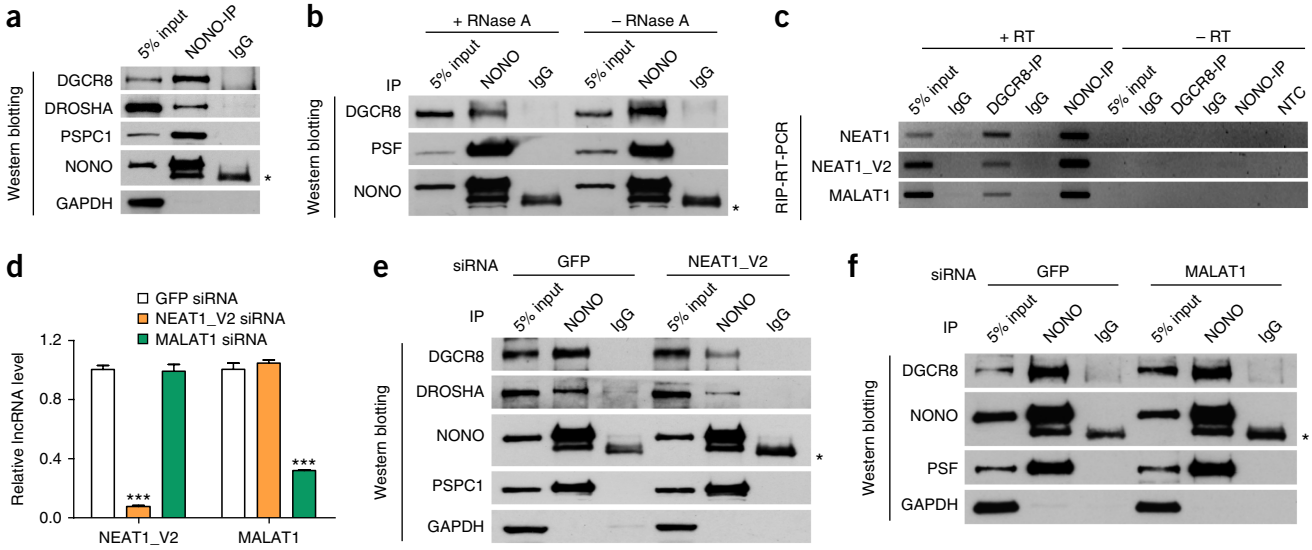

Figure 4 NEAT1 bridges NONO–PSF and the Microprocessor. (**a**) Co-IP of endogenous NONO with PSPC1 and the Microprocessor DROSHA–DGCR8. (**b**) RNA-dependent interactions between NONO–PSF and the Microprocessor. (**c**) RT-PCR analysis of NEAT1_V1, NEAT1_V2, and MALAT1 in NONO and DGCR8 immunoprecipitates. (**d**) Knockdown efficiency of NEAT1_V2 and MALAT1, quantified by RT-qPCR and normalized against *GAPDH* mRNA. Bar graphs represent mean ± s.e.m. ($n = 3$ technical replicates). $^{***}P < 0.001$, determined by two-tailed Student's t test. (**e**,**f**) Western blot analysis of NONO and Microprocessor interactions in response to knockdown of NEAT1_V2 (**e**) or MALAT1 (**f**). Single asterisk indicates IgG heavy chain. Uncropped images of western blots are shown in **Supplementary Data Set 1**. Source data for graphs are available online.

of which was crosslinked to RNA, as detected by ^{32}P labeling with T4 polynucleotide kinase (**Fig. 3a** and **Supplementary Fig. 5a**). We separately isolated protein–RNA adducts after trimming RNA with micrococcal nuclease for CLIP-seq library construction. Libraries from reciprocal immunoprecipitation showed high reproducibility among all NONO–PSF CLIP-seq experiments (**Supplementary Fig. 5b**). We thus combined uniquely mapped, PCR-duplicate-removed reads, obtaining ~14 million reads for NONO and ~18 million reads for PSF (sequencing statistics in **Supplementary Table 4**). The deduced NONO and PSF binding peaks were similarly distributed in the human genome with a large fraction on intronic and 3′ UTR regions (**Supplementary Fig. 5c**), consistent with their established roles in pre-mRNA processing[32,34,42].

Importantly, we found that both NONO and PSF bound 263 transcribed pri-miRNAs, about two-thirds of expressed pri-miRNAs in HeLa cells (**Fig. 3b**,**c**), as illustrated by their highly discrete binding on all six pri-miRNAs encoded in the *pri-miR-17-92a* locus (**Fig. 3d**), as well as on many other representative pri-miRNAs (**Supplementary Fig. 5d**). We noted that such PSF binding on expressed pri-miRNA was not as prevalent in HepG2 as we observed in HeLa cells, based on a similar analysis of the existing PSF eCLIP data from the ENCODE consortium (data not shown). Additionally, the number of downregulated miRNAs was clearly larger than that of NONO–PSF-bound pri-miRNAs. These observations imply potential cell-type specificity with respect of NONO–PSF binding on expressed pri-miRNAs and suggest roles of other paraspeckle-associated RBPs in pri-miRNA processing, which were likely affected by induced paraspeckle disassembly[28].

Interestingly, we also detected prevalent binding of the NONO–PSF heterodimer on NEAT1 but with a dramatically distinct binding pattern compared to their discrete binding on pri-miRNAs (**Fig. 3e**). Such continuous binding on both ends of NEAT1 is consistent with the proposed structure of paraspeckles having both of its ends exposed at the periphery of this subnuclear domain[43]. For comparison, we also displayed the published DGCR8 CLIP-seq data on NEAT1 (ref. 44). Although the DGCR8 CLIP-seq read density is relatively low on the lncRNA, the reads showed two binding clusters: one at the 5′ end and the other on pri-miR-612 at the 3′ end of NEAT1 (marked at the bottom of **Fig. 3e**). Notably, although NEAT1 is extremely abundant in the cell, we found that mature miR-612 was nearly undetectable from our small RNA-seq experiments or by RT-qPCR (data not shown). We further confirmed its poor processing by comparing a pri-miR-612-processing reporter with the reporter derived from pri-miR-17-92a in response to Microprocessor knockdown (**Supplementary Fig. 6a**). This finding suggested that miR-612 might be a 'pseudo' miRNA, and its primary function might be to serve as an anchor for attracting the Microprocessor to NEAT1. These data therefore begin to paint a general picture in which NEAT1 might function as a scaffold, not only for a large number of RBPs, but also for the Microprocessor, thereby facilitating their kinetic interactions that lead to more efficient pri-miRNA processing in the nucleus.

NEAT1 mediates interaction of NONO–PSF with the Microprocessor

To provide evidence for NEAT1-mediated interactions between NONO–PSF and the Microprocessor, we performed reciprocal immunoprecipitation with NONO and DGCR8, and found that NONO was indeed able to bring down both endogenous and exogenous FLAG-tagged DGCR8 in HeLa cells (**Fig. 4a** and **Supplementary Fig. 6b**,**c**). Importantly, RNase A treatment greatly reduced the interactions of NONO with DGCR8, but not with PSF, indicating direct protein–protein interactions between NONO and PSF and RNA-mediated interactions between NONO–PSF and the Microprocessor (**Fig. 4b**). As predicated, specific antibodies against DGCR8 and NONO also brought down both isoforms of NEAT1 (**Fig. 4c**). Interestingly, both also pulled down MALAT1, which is known to interact with numerous RBPs involved in pre-mRNA splicing[37]. To determine whether any of these lncRNAs mediated the interactions between NONO–PSF and the Microprocessor, we performed siRNA knockdown (**Fig. 4d**) and found that NEAT1 knockdown largely abolished the *in vivo* interactions

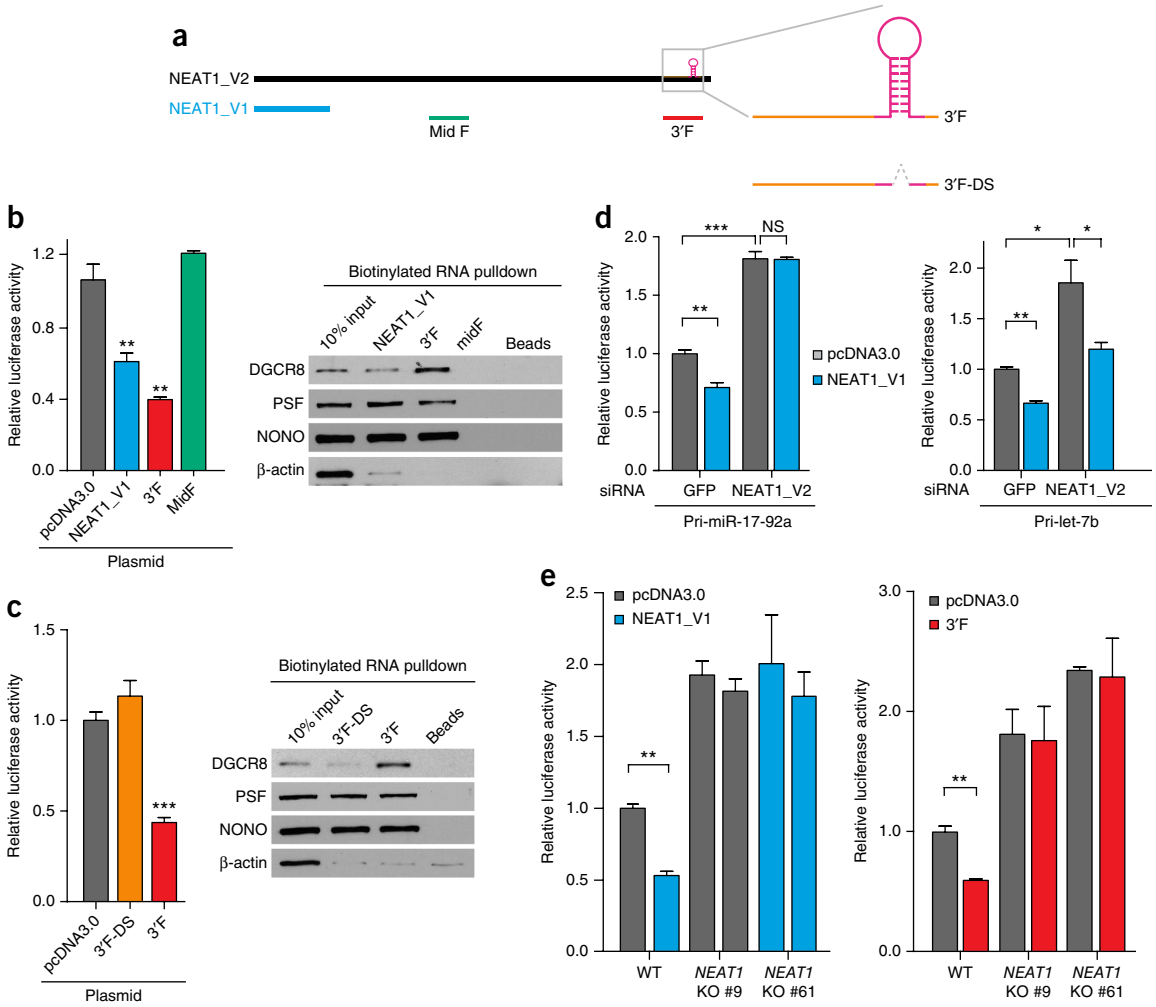

Figure 5 NEAT1-mediated interaction networks for enhancing pri-miRNA processing. (**a**) Illustration of NEAT1_V1, NEAT1_V2 and derived RNA fragments. Highlighted on the right are pri-miR-612 near the 3′ end of NEAT1_V2 and the 3′ fragments before (3′F) or after deletion of the pre-miR-612 stem-loop (3′F-DS). (**b**,**c**) Enhanced processing of the pri-miR-17-92a reporter by NEAT_V1 and 3′F, but not a middle fragment from NEAT1_V2 (midF) or 3′F-DS (left). Right panels show RNA pulldowns from HeLa nuclear extracts, analyzed by western blot for NONO–PSF and DGCR8. (**d**) KD of NEAT1_V2 diminished the enhancement of pri-miRNA processing by overexpressed NEAT_V1 on the pri-miR-17-92a (left) or pri-Let-7b (right) processing reporter. (**e**) Knockout of *NEAT1*_V2 prevented the enhancement of pri-miRNA processing by overexpressed NEAT_V1 and the 3′F on the pri-miR-17-92a processing reporter. Bar graphs in **b**–**e** show mean ± s.e.m. ($n = 3$ independent cell cultures). pcDNA3.0 is an empty plasmid control. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant, determined by two-tailed Student's t test. Uncropped images of western blots in **b** and **c** are shown in **Supplementary Data Set 1**. Source data for graphs are available online.

of NONO with the Microprocessor without affecting its interaction with PSPC1 (**Fig. 4e**). By contrast, MALAT1 knockdown showed no impact on any of these interactions (**Fig. 4f**). These data strongly suggested that NEAT1 specifically bridged the interactions between paraspeckle components and the Microprocessor in the cell.

NEAT1_V1 enhances pri-miRNA processing in a NEAT1_V2 dependent manner

Because the NONO–PSF heterodimer interacts with numerous regions in NEAT1 and the smaller V1 isoform of NEAT1 has been shown to enhance paraspeckle formation[28,45], we next determined whether V1 and some representative fragments from full-length NEAT1 (**Fig. 5a**) were able to enhance pri-miRNA processing. Using the pri-miR-17-92a-processing reporter (**Fig. 2a**), we found that transfected V1 was indeed stimulatory to pri-miR-17-92a processing, and so was a V2 fragment from the 3′ end (3′F), but not a middle fragment (midF) (**Fig. 5b**, left). When the stem–loop of pri-miR-612 in the 3′F was deleted (3′F-DS), the enhancement effect was lost (**Fig. 5c**, left). To further explore the molecular basis for enhanced pri-miR-17-92a processing, we incubated nuclear extracts with various *in vitro*–transcribed RNAs to determine their abilities to bridge the interactions between NONO–PSF and the Microprocessor. We found that both V1 and the 3′F, but not 3′F-DS, were able to efficiently bring down NONO–PSF and DGCR8 (**Fig. 5b**,**c**, right).

Because an earlier observation indicated that V1 was able to enhance the appearance of paraspeckles, but only in the presence of full-length NEAT1 (refs. 28,45), we next tested whether full-length NEAT1 was required for enhanced pri-miR-17-92a processing by V1. We observed that the enhancement was lost in NEAT1-knockdown cells (**Fig. 5d**, left). We made a similar observation using another Let-7b–based pri-miRNA-processing reporter, although V1 continued to show some effect on this reporter in NEAT1-depleted cells (**Fig. 5d**, right). These observations suggest a broad effect of full-length NEAT1. To further confirm this finding, we took advantage of

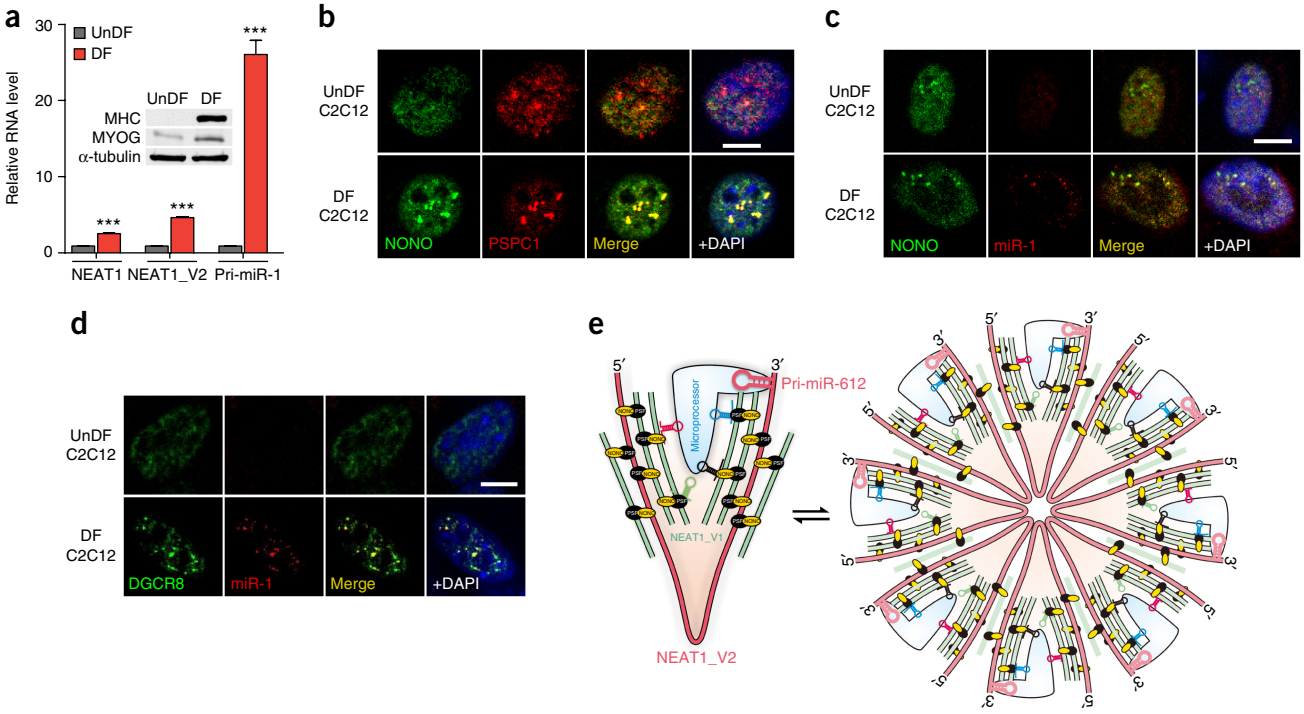

Figure 6 Localization of induced pri-miR-1 in paraspeckles in differentiated C2C12 cells and the proposed bird nest model. (**a**) Expression levels of NEAT1 and pri-miR-1 quantified by RT-qPCR in undifferentiated (unDF) and differentiated (DF) C2C12 cells. Inset shows the induction of the differentiation markers MHC and MYOG by western blot. Bar graphs represent mean ± s.e.m. ($n = 3$ technical replicates). ***$P < 0.001$; determined by two-tailed Student's t test. Uncropped images of western blots are shown in **Supplementary Data Set 1**. Source data for the bar graphs are available online. (**b**) Enhanced paraspeckles after C2C12 differentiation, detected by NONO and PSPC1 immunostaining. (**c**) FISH analysis of inducible pri-miR-1. No pri-miR-1 signal was detectable in C2C12 cells before differentiation. In differentiated C2C12 cells, induced pri-miR-1 colocalizes with NONO in paraspeckles. (**d**) Colocalization of induced pri-miR-1 with DGCR8 in differentiated C2C12 cells. Scale bars in **b**–**d** indicate 10 μm. (**e**) Proposed 'bird nest' model for NEAT1-orchestrated enhancement of pri-miRNA processing. Multiple RBPs, including the NONO–PSF heterodimer, interact with NEAT1_V2. Additional NEAT1_V1 and NEAT1_V2 can be recruited to build a bird nest–like structure. Distinct RBPs may bring different pri-miRNAs to the nest and various RNA secondary structures in NEAT1, including a poorly processed pri-miR-612 near the 3′ end of NEAT1_V2, may help recruit Microprocessor. These NEAT1-containing RNPs may exist in both the microscopic form (left) or become 'aggregated' to generate larger structures visible as paraspeckles (right). In both forms, these RNA-orchestrated structures may create proximity between pri-miRNAs and the Microprocessor to enhance the kinetics of pri-miRNA processing.

NEAT1-null cell lines we generated (**Fig. 1d**) and tested the requirement of full-length NEAT1 for stimulated pri-miRNA processing by the small V1 isoform or the 3′F of V2. Consistent with results from the NEAT1-knockdown experiments, we observed that both overexpressed V1 and the 3′F were able to enhance pri-miRNA processing in wild-type but not *NEAT1*-null cells (**Fig. 5e**). Combined, these data suggest that full-length NEAT1 functions in providing a platform for enhanced pri-miRNA processing in the nucleus.

Involvement of paraspeckles in pri-miRNA processing

A recent study suggested two populations of NEAT1-containing ribonucleoprotein particles (RNPs) in mammalian cells, one in the form of numerous microscopic structures throughout the nucleus and the other as paraspeckles[46], the latter of which may be the aggregated form of the former. In literature, the data on the involvement of paraspeckles have been controversy at this point. On one hand, it has been demonstrated that at least a fraction of pri-miRNAs are processed co-transcriptionally[9], and retarded release of pri-miRNAs from chromatin appears to be important for their efficient processing[7,47]. Accordingly, DGCR8 has been localized in a largely diffused pattern in the nucleoplasm[7]. These data suggest that paraspeckles visible under the microscope may not correspond to cellular locations for pri-miRNA processing. On the other hand, one report indicated that the FLAG-tagged exogenous DGCR8 was localized adjacent to nuclear speckles[48], and certain induced pri-miRNAs (for example, pri-miR-155) also became localized near nuclear speckles[7]. However, none of these studies verified adjacent nuclear speckles as paraspeckles by costaining with a paraspeckle marker. Therefore, although inconclusive, these existing data suggest that a subset of pri-miRNAs, especially those that are highly induced, as well as a fraction of the Microprocessor were detectable in a localized fashion under certain experimental conditions.

Given such controversy, we sought to localize DGCR8 by immunochemistry and pri-miRNAs by fluorescence *in situ* hybridization (FISH) under various conditions, but we rarely detected signals in paraspeckles. Reasoning that such localization might become detectable only with highly expressed pri-miRNAs, we took advantage of the dramatic induction of the *pri-mir-1* gene in C2C12 cells upon differentiation[49], which we confirmed. We detected an increase in *NEAT1* expression in differentiated C2C12 cells (**Fig. 6a**). By immunostaining with PSPC1, we found that undifferentiated C2C12 myoblasts exhibited detectable paraspeckles, but the appearance of paraspeckles became much stronger after C2C12 differentiation into myotubes, consistent with the induction of *NEAT1* (**Fig. 6b**). We next performed FISH for pri-miR-1 and found that induced pri-miR-1 was indeed detectable in multiple foci and colocalized with NONO

only on differentiated C2C12 cells (**Fig. 6c**). Under these conditions, we also observed that induced pri-miR-1 colocalized with DGCR8 (**Fig. 6d**). It is important to disclose that we observed such DGCR8 foci under FISH conditions but rarely saw such foci under standard immunostaining conditions, even on differentiated C2C12 cells (data not shown). These observations imply that DGCR8 may be engaged in other RNA metabolism pathways in the nucleoplasm, as suggested by its broad RNA-binding profile[44], which may mask its localization in paraspeckles. Under FISH conditions, some of the DGCR8 interactions were weakened in the nucleoplasm, despite hybridization on fixed cells, whereas interactions in paraspeckles were preserved, thus providing a plausible explanation to the controversy in the field. Importantly, our data now suggest the involvement of NEAT1-containing RNPs in the global modulation of pri-miRNA processing, either in the form of microscopic structures throughout the nucleus or in a more aggregated form in nuclear paraspeckles, similar to cotranscriptional and post-transcriptional pre-mRNA splicing with respect to nuclear speckles.

DISCUSSION

Considering all data presented in this study, we propose a bird nest model for NEAT1-mediated enhancement of pri-miRNA processing (**Fig. 6e**, left). The lncRNA NEAT1 may provide a scaffold for the NONO–PSF heterodimer and many other RBPs[28]. These RBPs may bind additional NEAT1_V1 and NEAT1_V2 isoforms as well as pri-miRNAs to form a bird-nest-like structure in the nucleus. As NEAT1 likely contains various hairpin structures that resemble pri-miRNAs, as exemplified by pri-miR-612, those secondary structures may also help attract the Microprocessor, thus facilitating kinetic interactions between pri-miRNAs and their processing machinery. These microscopic NEAT1-containing RBPs may become further aggregated to give rise to the appearance of paraspeckles (**Fig. 6e**, right), especially after cell differentiation, during which some specific components of paraspeckles are induced[35].

Although a role, or roles, of paraspeckles visible under a microscope in certain aspects of regulated gene expression will continue to be a subject of debate, our current data provide evidence for its active participation in post-transcriptional pri-miRNA processing and perhaps other RNA metabolism activities. It is well known in the field that DGCR8 is largely distributed in a diffused pattern in the nucleus, which likely reflects its involvement in multiple RNA metabolism pathways, consistent with its limited interaction with expressed pri-miRNAs and NEAT1 (**Fig. 3e**, bottom) based on the published DGCR8 RNA-binding profile[44]. We found that under certain conditions, such as on cells treated for FISH, a fraction of DGCR8 became detectable on paraspeckles where it colocalized with highly induced pri-miR-1 in differentiated C2C12 cells. This observation implies that the localization of DGCR8 in paraspeckles might be largely masked in most cell types. Because ~80% of pri-miRNAs reside in introns of pre-mRNAs[4,8] and the structural integrity of paraspeckle depends on on-going transcription[12,45], it is tempting to speculate that its spatial relationship with splicing-factor-enriched speckles might result from the coprocessing of certain pri-miRNAs and pre-mRNAs in the nucleus.

With respect to paraspeckle-associated RBPs, we noted that NONO–PSF binding on pri-miRNAs were more prevalent in HeLa cells than in HepG2 cells, implying a degree of cell-type specificity in terms of divided labors of different paraspeckle-associated RBPs in pri-miRNA processing. It is also curious that the paraspeckle-associated protein PSPC1 appears to suppress pri-miRNA processing. This might be related to the observation that PSPC1 belongs to a distinct family of paraspeckle-associated RBPs that are not essential for paraspeckle formation or maintenance, which requires further investigation on its regulatory mechanism.

Last but not least, we need to consider the fact that *Neat1*-null mouse does not have gross phenotype[18], suggesting that NEAT1-enhanced pri-miRNA processing may not be essential for cell survival; however, such a process may still critically contribute to specific gene expression programs under certain developmental and/or pathological conditions, as seen in *Neat1*-null animals[26,35,50–54]. In any case, the data presented in this report suggest a potential new role of NEAT1-containing RNPs, either in its microscopic form or as part of paraspeckles, which provides a new angle to envision and investigate the biological function of this intriguing nuclear subdomain in diverse developmental and disease processes.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by grants from the National Key R&D Program of China (2017YFA0504400), the 111 Program of China (B06018), NIH (HG004659, GM049369, GM052872), the National Natural Science Foundation of China (31000573), and the Postdoctoral Science Foundation of China (20090451074).

AUTHOR CONTRIBUTIONS

Q.-J.W., L.J., C.S., and X.-D.F. designed the experiments. L.J., Q.-J.W., and C.S. performed most experiments. J.Z., B.Y., and Y. Zhou analyzed the data; G.C., H.L., L.-T.G., Y. Zhang, Y.W. and G.W.Y. contributed to sequencing, cell lines, and data interpretation. L.J., C.S., Y. Zhou and X.-D.F. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
2. Voinnet, O. Origin, biogenesis, and activity of plant microRNAs. *Cell* **136**, 669–687 (2009).
3. Pasquinelli, A.E. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev. Genet.* **13**, 271–282. http://dx.doi.org/10.1038/nrg3162 (2012).
4. Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **14**, 10A, 1902–1910 (2004).
5. Ha, M. & Kim, V.N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).
6. Liu, H. *et al.* HP1BP3, a chromatin retention factor for co-transcriptional microRNA processing. *Mol. Cell* **63**, 420–432 (2016).
7. Pawlicki, J.M. & Steitz, J.A. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J. Cell Biol.* **182**, 61–76 (2008).
8. Kim, Y.K. & Kim, V.N. Processing of intronic microRNAs. *EMBO J.* **26**, 775–783 (2007).
9. Morlando, M. *et al.* Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Biol.* **15**, 902–909 (2008).
10. Ballarino, M. *et al.* Coupled RNA processing and transcription of intergenic primary microRNAs. *Mol. Cell. Biol.* **29**, 5632–5638 (2009).
11. Fox, A.H. *et al.* Paraspeckles: a novel nuclear domain. *Curr. Biol.* **12**, 13–25 (2002).
12. Fox, A.H. & Lamond, A.I. Paraspeckles. *Cold Spring Harb. Perspect. Biol.* **2**, a000687 (2010).
13. Huang, S. & Spector, D.L. Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. *Genes Dev.* **5**, 12A, 2288–2302 (1991).
14. Hall, L.L., Smith, K.P., Byron, M. & Lawrence, J.B. Molecular anatomy of a speckle. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **288**, 664–675 (2006).

15. Brown, J.M. *et al.* Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J. Cell Biol.* **182**, 1083–1097 (2008).
16. Yamazaki, T. & Hirose, T. The building process of the functional paraspeckle with long non-coding RNAs. *Front. Biosci. (Elite Ed.)* **7**, 1–41 (2015).
17. Nakagawa, S. *et al.* Malat1 is not an essential component of nuclear speckles in mice. *RNA* **18**, 1487–1499 (2012).
18. Nakagawa, S., Naganuma, T., Shioi, G. & Hirose, T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J. Cell Biol.* **193**, 31–39 (2011).
19. Hutchinson, J.N. *et al.* A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**, 39 (2007).
20. Wilusz, J.E., Freier, S.M. & Spector, D.L. 3′ end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135**, 919–932 (2008).
21. Clemson, C.M. *et al.* An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **33**, 717–726 (2009).
22. Sasaki, Y.T., Ideue, T., Sano, M., Mituyama, T. & Hirose, T. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. USA* **106**, 2525–2530 (2009).
23. Prasanth, K.V. *et al.* Regulating gene expression through RNA nuclear retention. *Cell* **123**, 249–263 (2005).
24. Chen, L.L. & Carmichael, G.G. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell* **35**, 467–478 (2009).
25. Hirose, T. *et al.* NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol. Biol. Cell* **25**, 169–183 (2014).
26. Imamura, K. *et al.* Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol. Cell* **53**, 393–406 (2014).
27. West, J.A. *et al.* The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* **55**, 791–802 (2014).
28. Naganuma, T. *et al.* Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* **31**, 4020–4034 (2012).
29. Urcuqui-Inchima, S., Castaño, M.E., Hernandez-Verdun, D., St-Laurent, G. III & Kumar, A. Nuclear Factor 90, a cellular dsRNA binding protein inhibits the HIV Rev-export function. *Retrovirology* **3**, 83 (2006).
30. Guil, S. & Cáceres, J.F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.* **14**, 591–596 (2007).
31. Chen, L.L. & Carmichael, G.G. Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats. *Cell Cycle* **7**, 3294–3301 (2008).
32. Kaneko, S., Rozenblatt-Rosen, O., Meyerson, M. & Manley, J.L. The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3′ processing and transcription termination. *Genes Dev.* **21**, 1779–1789 (2007).
33. Lowery, L.A., Rubin, J. & Sive, H. Whitesnake/sfpq is required for cell survival and neuronal development in the zebrafish. *Dev. Dyn.* **236**, 1347–1357 (2007).
34. Liang, S. & Lutz, C.S. p54nrb is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. *RNA* **12**, 111–121 (2006).
35. Sunwoo, H. *et al.* MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* **19**, 347–359 (2009).
36. Tripathi, V. *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **39**, 925–938 (2010).
37. Spector, D.L. & Lamond, A.I. Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* **3**, a000646 (2011).
38. Naganuma, T. & Hirose, T. Paraspeckle formation during the biogenesis of long non-coding RNAs. *RNA Biol.* **10**, 456–461 (2013).
39. Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P. & Blelloch, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* **22**, 2773–2785 (2008).
40. Miyamoto, K., Sakurai, H. & Sugiura, T. Proteomic identification of a PSF/p54nrb heterodimer as RNF43 oncoprotein-interacting proteins. *Proteomics* **8**, 2907–2910 (2008).
41. Passon, D.M. *et al.* Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. *Proc. Natl. Acad. Sci. USA* **109**, 4846–4850 (2012).
42. Patton, J.G., Porro, E.B., Galceran, J., Tempst, P. & Nadal-Ginard, B. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. *Genes Dev.* **7**, 393–406 (1993).
43. Souquere, S., Beauclair, G., Harper, F., Fox, A. & Pierron, G. Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol. Biol. Cell* **21**, 4020–4027 (2010).
44. Macias, S. *et al.* DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. *Nat. Struct. Mol. Biol.* **19**, 760–766 (2012).
45. Mao, Y.S., Sunwoo, H., Zhang, B. & Spector, D.L. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* **13**, 95–101 (2011).
46. Li, R., Harvey, A.R., Hodgetts, S.I. & Fox, A.H. Functional dissection of NEAT1 using genome editing reveals substantial localization of the NEAT1_1 isoform outside paraspeckles. *RNA* **23**, 872–881 (2017).
47. Pawlicki, J.M. & Steitz, J.A. Subnuclear compartmentalization of transiently expressed polyadenylated pri-microRNAs: processing at transcription sites or accumulation in SC35 foci. *Cell Cycle* **8**, 345–356 (2009).
48. Shiohama, A., Sasaki, T., Noda, S., Minoshima, S. & Shimizu, N. Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins. *Exp. Cell Res.* **313**, 4196–4207 (2007).
49. Zhang, X. *et al.* MicroRNA directly enhances mitochondrial translation during muscle differentiation. *Cell* **158**, 607–619 (2014).
50. Zeng, C. *et al.* Inhibition of long non-coding RNA NEAT1 impairs myeloid differentiation in acute promyelocytic leukemia cells. *BMC Cancer* **14**, 693 (2014).
51. Choudhry, H. & Mole, D.R. Hypoxic regulation of the noncoding genome and NEAT1. *Brief. Funct. Genomics* **15**, 174–185 (2016).
52. Choudhry, H. *et al.* Tumor hypoxia induces nuclear paraspeckle formation through HIF-2α dependent transcriptional activation of NEAT1 leading to cancer cell survival. *Oncogene* **34**, 4482–4490 (2015).
53. Nakagawa, S. *et al.* The lncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. *Development* **141**, 4618–4627 (2014).
54. Adriaens, C. *et al.* p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nat. Med.* **22**, 861–868 (2016).

ONLINE METHODS

Cell culture, plasmids, transfection, RIP-PCR and RT-qPCR. HeLa and C2C12 cells from ATCC were grown in DMEM (Gibco) supplemented with 10% FBS (FBS) plus 100 U penicillin/streptomycin (Gibco) at 37 °C in a humidified incubator with 5% CO_2. Both cultured HeLa and C2C12 cells were determined to be free from mycoplasma contamination.

All expression plasmids were cloned in pcDNA3.0 and luciferase reporter plasmids were cloned in psiCHECK2 between XholI and NotI sites. The siRNA-resistant FLAG-PSF and HA-NONO expression plasmids were generated by a PCR-based method (KOD Plus from TOYOBO) with specific primers containing site-specific mutations, listed in **Supplementary Table 1**.

Plasmids and siRNAs were transfected into cells with Lipofectamine2000 and RNAi Max, respectively, according to the manufacturer's protocols. Cells were harvested 48–72 h post-transfection for subsequent analysis. For double transfection, cells were first transfected with siRNA for 12 h and then with plasmid for another 24 h. Individual siRNA sequences are listed in **Supplementary Table 1**. Dual luciferase assays were performed 48 h post-transfection.

For RIP-PCR, RNA from immunoprecipitant was extracted with Trizol (ThermoFisher) and reverse transcribed with M-MLV (Promega) and random hexamers at 37 °C for 1 h. Quantitative analysis of miRNAs was performed with the Qiagen miScript II RT Kit, and real-time PCR was conducted by using a SYBR green master mix and gene-specific primers listed in **Supplementary Table 1**.

Immunoprecipitation, RNA pulldown, and western blotting. For immunoprecipitation, 2 μg specific antibody was coupled to Dynabeads at 4 °C for 2 h in 200 μl lysis buffer (50 mM Tris-HCl, pH8.0, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 1 mM PMSF and proteinase inhibitor cocktail) and washed three times with lysis buffer, then 300 μl of whole cell extracts was added to the beads. The mix was incubated with rotation at 4 °C for 2 h. For RNase A treatment, the immunoprecipitant was incubated in lysis buffer containing 200 μg/ml RNase A (Fermentas) at 37 °C for 10 min. The beads were washed four times with lysis buffer, resuspended in 1× SDS loading buffer, and boiled for analysis by SDS-PAGE. Western blotting was performed with standard protocol using specific antibodies listed in **Supplementary Table 2**.

For the preparation of nuclear extracts, 1×10^7 HeLa cells grown in 10-cm dishes were washed twice with 10 ml ice-cold PBS and collected by centrifugation at 2,000 r.p.m. for 5 min at 4 °C. Pelleted cells were resuspended in 1 ml CE-I buffer (10 mM HEPES, pH7.4, 60 mM KCl, 1 mM EDTA, 0.075% NP-40, 1 mM DTT) and incubated on ice for 5 min. Nuclei were isolated by centrifugation at 1,500 r.p.m. at 4 °C for 5 min, washed twice with 0.5 ml CE-II buffer (10 mM HEPES, pH7.4, 60 mM KCl, 1 mM EDTA, 1 mM DTT), resuspended in 0.5 ml NE buffer (20 mM HEPES, pH7.4, 420 mM KCl, 4 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 15% glycerol) by a briefly vortexing, then incubated with rotation at 4 °C for 60 min.

For RNA pulldown, biotin-labeled RNAs were prepared with the Biotin RNA labeling mix (Sigma) and T7 RNA polymerase (ThermoFisher). The reaction was carried out at 37 °C for 2 h, treated with 2 μl DNase I (Promega) at 37 °C for 15 min and then stopped by the addition of 2 μl 0.5 M EDTA (pH8.0). The resulting RNA was purified on a Micro Bio-Spin 30 Column (Bio-Rad) and stored in TE buffer (10 mM Tris-HCl, pH7.4). ~3 μg of biotin-labeled RNA was heated to 90 °C for 2 min, chilled on ice for 2 min, and incubated in 50 μl RNA structure buffer (10 mM Tris-HCl, pH 7.0, 0.1 M KCl, 10 mM $MgCl_2$) at room temperature for 20 min to allow RNA folding. Folded RNA was mixed with ~1 mg of nuclear extracts diluted in RIP buffer (25 mM Tris-HCl, pH 7.0, 150 mM KCl, 0.5 mM DTT, 0.5% NP-40, 1 mM PMSF and protease inhibitor cocktail) and incubated at 4 °C for 1 h followed by the addition of 30 μl of streptavidin beads that had been prewashed twice with RIP buffer. The reaction was further incubated for another 1 h, washed three times with cold RIP buffer, and boiled in 30 μl 1× SDS loading buffer for analysis by western blotting.

Global miRNA profiling and CLIP-seq. Isolated total RNA (5–8 μg) was mixed with 10 μM pre-adenylated 3′ linker in 1× RNA ligase buffer, containing 50% (w/v) PEG 8000, 1 μl RNase inhibitor, and 1 μl T4 RNA ligase 2 (NEB), and incubated at room temperature for 1 h, then at 16 °C for 3 h. After 3′ linker ligation, the reaction was size fractionated on a 15% polyacrylamide gel and RNAs between 42 to 54 nt were recovered by incubating the corresponding gel slice at 4 °C in 600 μl of 0.3 M NaCl with constant agitation. 5′ linker ligation was performed next in 1× RNA ligase buffer containing 50% (w/v) PEG 8000, 1 mM ATP, 10 μM 5′ linker, 1 μl RNase inhibitor and 1 μl T4 RNA ligase 1 (NEB) at 37 °C for 4 h. Linker-ligated RNAs were reverse transcribed with Superscript III (Life Technology) and PCR-amplified with Phusion polymerase (ThermoFisher). The product was fractionated on a 4% agarose gel, and recovered DNA was quantified for deep sequencing.

Duplicated experiments were performed for treatment condition and multiplexed sequenced in one Illumina Hi-seq 4000 lane (sequencing statistics in **Supplementary Table 4**). Reads were decoded by index sequences without mismatch, and 1–40-nt target sequences were saved for downstream analysis. The 3′ adaptor of reads (NNCTCGTATGCCGTCTTCTGCTTG) was first trimmed by using cutadapt program[55] with parameters '-O 3 -e 0.25 -q 20'. The 5′ random index (NNNTC) was then removed and only reads of ≥16 nt were kept for mapping. The annotated miRNAs in miRBase version 20 (ref. 56) were used as a reference. The reads were mapped to the human genome (hg19) using the Bowtie program[57], and the expression of mature miRNAs were quantified with program miRDeep2 (ref. 58). The counts for mature miRNAs from two replicates were combined, and the counts from different samples were normalized according to the published procedure[39], using the number of reads mapping to rRNAs, snoRNAs, tRNAs, and spike-in RNA (CTCAGGATGGCGGAGCGGTCT) as internal controls. A 1.5-fold change or larger of reads from knockdown samples relative to the siGFP sample were computed to identify differentially expressed miRNAs, as summarized in **Supplementary Table 4**.

CLIP-seq for NONO and PSF, as well as associated data analysis, were conducted as previously described[59,60], and the peak calling was done with CLIPper[59]. The distribution of binding was computed with DeepTools2 (ref. 61). The sequencing statistics are listed in **Supplementary Table 4**.

Genomic engineering by CRISPR–Cas. sgRNAs (**Supplementary Table 1**) were designed using the CRISPR tool at http://crispr.mit.edu. To knockout *PSPC1*, annealed DNA oligonucleotides were cloned into px459 at the BbsI restriction site. To knockout *NEAT1*, four pairs of sgRNA sequences (#1 and #3; #1 and #4; #2 and #3; #2 and #4) were cloned into the PiggyBac plasmid (PBC2), each under a separate U6 promoter. The PBC2 plasmids containing individual pairs of sgRNAs and the plasmid expressing Cas9 were cotransfected in HeLa cells using Lipofectamine2000. After 24 h post-transfection, *PSPC1* KO cells were selected with puromycin (Sigma) for 4 d, and *NEAT1* KO cells were selected with hygromycin (Roche) for 4 d. Live cells were cultured in fresh DMEM containing 10% FBS without antibiotics to allow recovery for 1–2 days before isolating single clones. For *PSPC1* KO cell lines, out of 48 clones obtained, two homozygous clones were picked for functional analysis. For *NEAT1* KO, out of 148 clones obtained, four homozygous clones (one from sgRNA pair #1 and #3, one from sgRNA pair # 2 and # 3, two from sgRNA # 2 and #4) were picked for verification by direct sequencing and subsequent functional studies.

RNA FISH on C2C12 cells. C2C12 cells grown to 80% confluence were induced to differentiate into myotubes by incubation in DMEM supplemented with 2% horse serum (Gibco) for two days. FISH was performed as previously described[22]. Fixed cells on a coverslip were dehydrated with 70%, 95%, and 100% ethanol for 5 min each and incubated with prehybridization buffer (2× SSC, 1× Denhardt's solution, 50% formamide, 10 mM EDTA, 100 μg/ml yeast tRNA, and 0.01% Tween 20) at 37 °C for 1 h. Prehybridized coverslips were incubated in hybridization buffer (5% dextran sulfate in prehybridization buffer plug Dig-labeled RNA Probe) for 16–18 h at 37 °C. Coverslips were washed twice with buffer A (2× SSC, 50% formamide, and 0.01% Tween 20) at 37 °C for 20 min, twice with buffer B (2× SSC and 0.01% Tween 20), and once with buffer C (0.1× SSC and 0.01% Tween 20). Coverslips were then blocked with blocking buffer (50 mM Tris, pH8.0, 100 mM NaCl, 0.1%TritonX-100, 3%NGS, 0.1%BSA) for 1 h and incubated with anti-DIG antibody (ThermoFisher). For colocalization with protein markers, cells were blocked with blocking buffer at 37 °C for 20 min and incubated with specific antibodies diluted in the blocking buffer for 1 h. After washing the cells three times with 1× TBST, fluorescent-labeled secondary antibodies (Alexa 594 anti-sheep or Alexa 488 anti-rabbit) were applied and the cells were incubated for 1 hr and then washed three times with 1× TBST and counterstained with DAPI before being mounted onto glass slides for microscopy. Images were taken

with a Leica SP8 microscope with a 63× objective lens. The primers used for probe preparation are listed in **Supplementary Table 1** and the antibodies used for immunostaining are listed in **Supplementary Table 2**.

Statistical analysis. Data shown for each experiment was based on 3 or 4 technical replicates, as indicated in individual figure legends. Data are presented as means ± s.e.m, and *P* values were determined by two-tailed Student's *t* test. All experiments were further confirmed by biological repeats.

A **Life Science Reporting Summary** for this paper is available.

Data availability. All deep sequencing data from this study have been deposited in the Gene Expression Omnibus (GEO) under series accession number GSE90650. Source data for the graphs in all figures are available online. Other data are available upon reasonable request.

55. Katz, Y., Wang, E.T., Airoldi, E.M. & Burge, C.B. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods* **7**, 1009–1015 (2010).
56. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39**, D152–D157 (2011).
57. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
58. Friedländer, M.R., Mackowiak, S.D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).
59. Lovci, M.T. *et al.* Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.* **20**, 1434–1442 (2013).
60. Shao, C. *et al.* Mechanisms for U2AF to define 3′ splice sites and regulate alternative splicing in the human genome. *Nat. Struct. Mol. Biol.* **21**, 997–1005 (2014).
61. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W1, W160–W165 (2016).

doi:10.1038/nsmb.3455

nature research

Corresponding author(s): Yu Zhou
Xiang-Dong Fu

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

In most RT-qPCR and luciferase assays, the sample size was based on standard triplicated experiments.

Data exclusions

Describe any data exclusions.

No data were excluded from the analyses.

Replication

Describe whether the experimental findings were reliably reproduced.

All of the experiments were confirmed by both technical and biological repeats as described in Methods.

Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

n/a

Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

n/a

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.3455

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	cutadapt; Bowtie; miRDeep2; CLIPper; DeepTools2; Excel; Graphpad Prism 5; Adobe Illustrator CC; Adobe Photoshop CC; Image J 1.5.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	We provided all critical information on antibodies in the Supplementary Table 2, including catalog number, host, supplier and specific dilution for each antibody we used in this study. All validation information for the antibodies can be searched in antibodypedia.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	HeLa and C2C12 cell lines were purchased from ATCC.
b. Describe the method of cell line authentication used.	The cell lines were not further authenticated.
c. Report whether the cell lines were tested for mycoplasma contamination.	We periodically checked potential contamination with mycoplasma, which causes retarded cell growth and low pH in the media. Mycoplasma was tested by Hoechst staining o f the cells according to Young L.et al., Nature Protocols,2010.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	HeLa and C2C12 cell lines used in this study are not listed in the database.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	The study did not involve human participants.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.3455