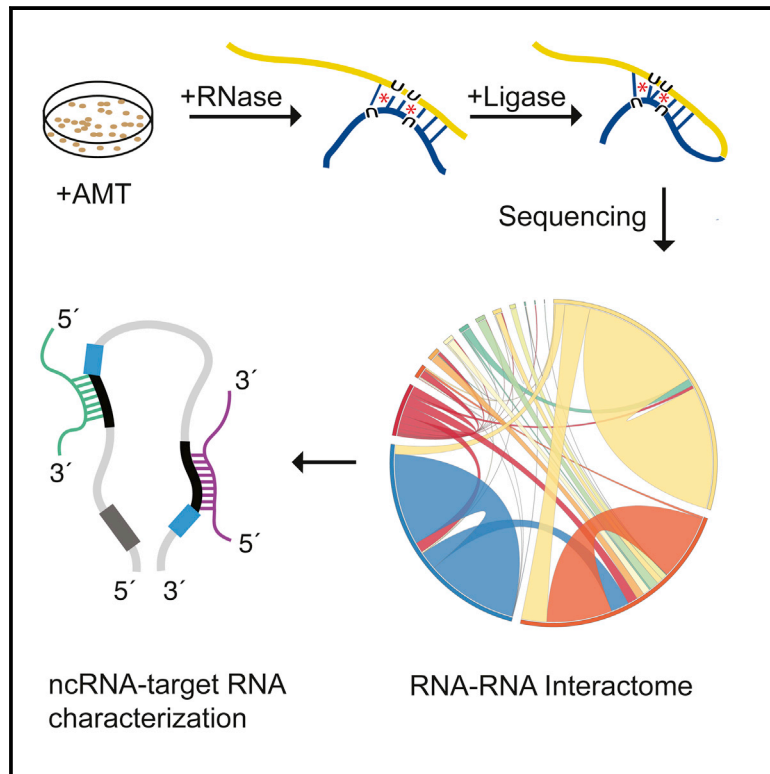# Molecular Cell

Global Mapping of Human RNA-RNA Interactions

Graphical Abstract

Authors

Eesha Sharma, Tim Sterne-Weiler,
Dave O'Hanlon, Benjamin J. Blencowe

Correspondence

b.blencowe@utoronto.ca

In Brief

Most non-coding RNAs lack known functions or else are poorly characterized. Sharma, Sterne-Weiler et al. describe "LIGation of interacting RNA and high-throughput sequencing" (LIGR-seq), a method for the global-scale mapping of RNA-RNA interactions in vivo. Mapping the RNA-RNA interactome in human cells reveals unexpected functions for small nucleolar RNAs.

Highlights

- LIGR-seq is a method for the global-scale mapping RNA-RNA interactions in vivo

- LIGR-seq data reveal a complex RNA-RNA interactome in human cells

- Hundreds of trans-interactions involving known and orphan ncRNAs are detected

- The orphan snoRNA SNORD83B regulates levels of its LIGR-seq-detected target mRNAs

Accession Numbers

GSE80167

CrossMark

CellPress

Global Mapping of Human RNA-RNA Interactions

Eesha Sharma,[1,2,3] Tim Sterne-Weiler,[1,3] Dave O'Hanlon,[1] and Benjamin J. Blencowe[1,2,*]
[1]Donnelly Centre
[2]Department of Molecular Genetics
University of Toronto, Toronto, ON M53 3E1, Canada
[3]Co-first author
*Correspondence: b.blencowe@utoronto.ca
http://dx.doi.org/10.1016/j.molcel.2016.04.030

SUMMARY

The majority of the human genome is transcribed into non-coding (nc)RNAs that lack known biological functions or else are only partially characterized. Numerous characterized ncRNAs function via base pairing with target RNA sequences to direct their biological activities, which include critical roles in RNA processing, modification, turnover, and translation. To define roles for ncRNAs, we have developed a method enabling the global-scale mapping of RNA-RNA duplexes crosslinked in vivo, "LIGation of interacting RNA followed by high-throughput sequencing" (LIGR-seq). Applying this method in human cells reveals a remarkable landscape of RNA-RNA interactions involving all major classes of ncRNA and mRNA. LIGR-seq data reveal unexpected interactions between small nucleolar (sno) RNAs and mRNAs, including those involving the orphan C/D box snoRNA, SNORD83B, that control steady-state levels of its target mRNAs. LIGR-seq thus represents a powerful approach for illuminating the functions of the myriad of uncharacterized RNAs that act via base-pairing interactions.

INTRODUCTION

The formation of duplexes and other types of structured RNA represents a critical feature of most steps in the gene expression pathway. Key examples include dynamic interactions involving small nuclear (sn)RNA-snRNA and snRNA-precursor messenger (pre-m)RNA during the assembly and disassembly of spliceosomes (Wahl et al., 2009), interactions between amino-acylated transfer RNAs (tRNAs) and mRNAs that dictate peptide bond formation during translation (Noller, 2006), interactions between small nucleolar RNAs (snoRNAs) and target RNAs that guide the addition of RNA modifications (Lui and Lowe, 2013), and interactions between ncRNAs and mRNAs that control transcript turnover and translation (He and Hannon, 2004). Despite these examples, an obstacle confronting the detection and characterization of new and functionally important RNA-RNA interactions is the lack of a method permitting the systematic, global-scale mapping of RNA duplexes in vivo.

Methods have been described for the global-scale mapping of single- and double-stranded regions in RNA that couple high-throughput sequencing and RNA footprinting strategies utilizing structure-sensitive chemicals (Ding et al., 2014; Rouskin et al., 2014; Spitale et al., 2015) or nucleases (Underwood et al., 2010; Wan et al., 2014). While effective for mapping local primary and RNA secondary structure in vivo and in vitro, these approaches are not suitable for the detection of long-range structural interactions, nor can they identify intermolecular RNA interactions. Detection of long-range and other tertiary RNA interactions has been made possible in some cases by structure-based methods including nuclear magnetic resonance, X-ray crystallography, and cryo-electron microscopy, yet these methods currently are not high throughput (Feigon, 2015). In contrast, computational methods have been developed for the large-scale prediction of intra- and inter-molecular RNA interactions, often by taking advantage of phylogenetic comparisons that score the conservation of RNA duplexes. For example, bioinformatic approaches have predicted microRNA target sites in mRNA 3' UTR sequences (Oulas et al., 2015), and snoRNA-guided sites of RNA modification in ribosomal (r) RNA and snRNAs (Kehr et al., 2011; Tafer et al., 2010). However, computational approaches vary considerably in their accuracy, ultimately require experimental validation, and in many cases are not effective for identifying RNA-RNA interactions involving short or poorly conserved duplexes, or for more complex types of short- and long-range interactions.

A significant advance in mapping long-range nucleic acid interactions was afforded by the development of "proximity ligation" methods, in which junction sequences formed by the ligation of two interacting sequences are detected by amplification and sequencing. Initial applications included methods such as "3C" and "Hi-C" for the mapping of native chromatin conformation in vivo (Belton and Dekker, 2015; Dekker et al., 2002). More recently, RNA proximity ligation following immunoprecipitation of complexes of interest has been used to map Argonaute-bound microRNA-mRNA interactions (Grosswendt et al., 2014; Helwak et al., 2013), snoRNP-bound snoRNA-rRNA interactions (Kudla et al., 2011), and structured RNA bound by the Staufen protein (Sugimoto et al., 2015). It has also been used to map short- and long-range intramolecular RNA interactions in abundant ncRNAs in yeast and human cells, without a prior crosslinking step (Ramani et al., 2015). Limitations of these approaches are that they require prior knowledge—and specific affinity purification of—an interacting protein partner, and protocols that do not employ crosslinking in vivo to stabilize RNA

duplexes are prone to high false-positive and -negative detection rates due to undesired spurious RNA associations, or loss of unstable interactions, respectively. Thus, to determine the extent to which different classes of RNAs interact with each other in cells, and to assign possible roles to the massive repertoires of functionally uncharacterized RNAs in cells, the development of new methods that enable the systematic mapping of RNA duplexes in vivo is required.

To address these challenges, we developed LIGation of interacting RNA followed by high-throughput sequencing (LIGR-seq), a rapid and versatile method for the global-scale detection of RNA-RNA interactions in vivo that does not require prior knowledge of RNAs forming interactions, or of proteins required for such interactions. In this study, we employed LIGR-seq to map intermolecular RNA-RNA interactions in human 293T cells. We demonstrate that LIGR-seq detects with high specificity known snRNA-snRNA interactions within spliceosomal snRNPs and active spliceosomes, as well as known rRNA-rRNA interactions within the ribosome. Our method further illuminates an extensive landscape of intermolecular interactions involving all major classes of RNA, comprising previously unknown ncRNA-ncRNA interactions, ncRNA-mRNA interactions, and mRNA-mRNA interactions. These data reveal RNA targets and functions for snoRNAs, including interactions involving the orphan snoRNA SNORD83B that promotes the turnover of its target mRNAs. LIGR-seq is thus an effective tool for providing global-scale maps of in vivo RNA-RNA interactions that serve as a basis for uncovering functional roles of previously uncharacterized ncRNAs.

RESULTS

LIGR-seq (Figure 1A) employs in vivo crosslinking of RNA duplexes using the modified psoralen derivative 4′-aminomethyltrioxalen (AMT), which intercalates into RNA duplexes and, upon 365 nm UV irradiation, generates inter-strand adducts between juxtaposed pyrimidine bases (Calvet and Pederson, 1979). Following cell lysis and a limited single-strand S1 endonuclease digest, free RNA overhangs adjacent to duplexes are ligated using circRNA ligase. This ligase catalyzes the ATP-dependent ligation of proximal RNA ends and has optimal activity at elevated temperatures that reduce RNA hybridization. RNase R, a $3′ \rightarrow 5′$ exoribonuclease that digests linear and structured RNAs (Vincent and Deutscher, 2006), is then used to digest uncrosslinked RNA, thereby enriching AMT-crosslinked duplexes (Figures 1A and S1). Following reversal of crosslinks using 254 nm irradiation, the RNA samples are subjected to high-throughput sequencing to detect chimeras formed by ligation. To assess background ligation artifacts, uncrosslinked ("–AMT") samples are prepared in parallel, and unligated samples are used to determine the relative expression of transcripts forming chimeras for downstream normalization and analysis.

A computational method ("*Aligater*") employing recursive chaining of local transcriptome alignments was developed to identify chimeric reads (Figure 1B). Following filtering steps to remove artifacts, chimeras were classified as intra- or inter-molecular interactions (Figure 1B; Figures S1C–S1F; Supplemental Experimental Procedures for details). Reads representing intramolecular interactions were enriched in –AMT versus +AMT samples, and intermolecular interactions were enriched in the +AMT versus –AMT sample (Figure S1F). A probabilistic model was developed to assess the significance of detected interactions, using observed over expected ratios of chimeric reads. The observed and expected values relate to the abundance of transcripts forming the chimeras as measured using the plus and no ligase samples, respectively (Figure 1B; Figure S1D; Supplemental Experimental Procedures for details). We observe ratios of chimeric reads reflecting significant interactions in the +AMT versus –AMT sample that increase as a function of the expression of the interacting RNAs (Figure 1C). From these data, we estimate upper bounds to the false discovery rate (FDR) within the range of <4.4% for highly expressed transcripts (>250 RPM, ±AMT = 22.5) and <25% for relatively low expressed genes (>10 RPM, ±AMT = 4). These estimates are based on the assumption that –AMT interactions are true negatives. However, since stable, true-positive interactions such as U4 snRNA+U6 snRNA and 5S rRNA+28S rRNA (see below) are detected in both the +AMT and –AMT samples (Table S1), we expect the true FDR is below these upper bounds. Comparing the log ratios of observed versus expected counts of chimeric reads (OE_{+AMT}) for statistically significant chimeras in the +AMT sample in the biological replicate samples, we observe consistent results for significant interactions (Figure 1D; Spearman *Rho* = 0.38, p < 8.247e-06). These data thus indicate that LIGR-seq results in the reproducible detection of chimeric sequences formed as a consequence of RNA-RNA interactions crosslinked by AMT in vivo.

LIGR-Seq Detects Known RNA-RNA Interactions with High Specificity

To assess the specificity of LIGR-seq, we analyzed all possible pairwise interactions between snRNAs that form major (U2-dependent) and minor (U12-dependent) spliceosomes. As expected, we detect a strong signal for chimeric products comprising U4+U6 snRNAs (OE_{+AMT} = 241.5, p < 1.94e-272), which stably interact and crosslink within the context of U4/U6 snRNP and U4/U6.U5 tri-snRNP complexes (Figure 2A) (Behrens and Lührmann, 1991; Rinke et al., 1985). Chimeras comprising U2+U6 snRNAs were also abundant (OE_{+AMT} = 84.4, p < 5.0e-324), consistent with the dynamic interactions between these snRNAs during the formation of spliceosomes (Hall and Padgett, 1996; Madhani and Guthrie, 1992; Sun and Manley, 1995; Wassarman and Steitz, 1993). We also observe strong enrichment (OE_{+AMT} = 8332.3, p < 7.48e-278) for chimeras representing known interactions between the minor spliceosomal snRNAs U4ATAC+U6ATAC (Tarn and Steitz, 1996). However, in contrast to detection of U2+U6 snRNA interactions, the analogous U12+U6ATAC minor spliceosomal interactions were not significantly enriched (Figure 2B). This may reflect the relative abundance of the two types of spliceosome, where introns that are substrates for the minor spliceosome represent less than one percent of those that are spliced by the major spliceoesome (Will and Lührmann, 2005). However, biases associated with specific steps of the LIGR-seq protocol, such as the efficiency of the AMT crosslinking and ligation steps, which can be influenced by the length, composition, and
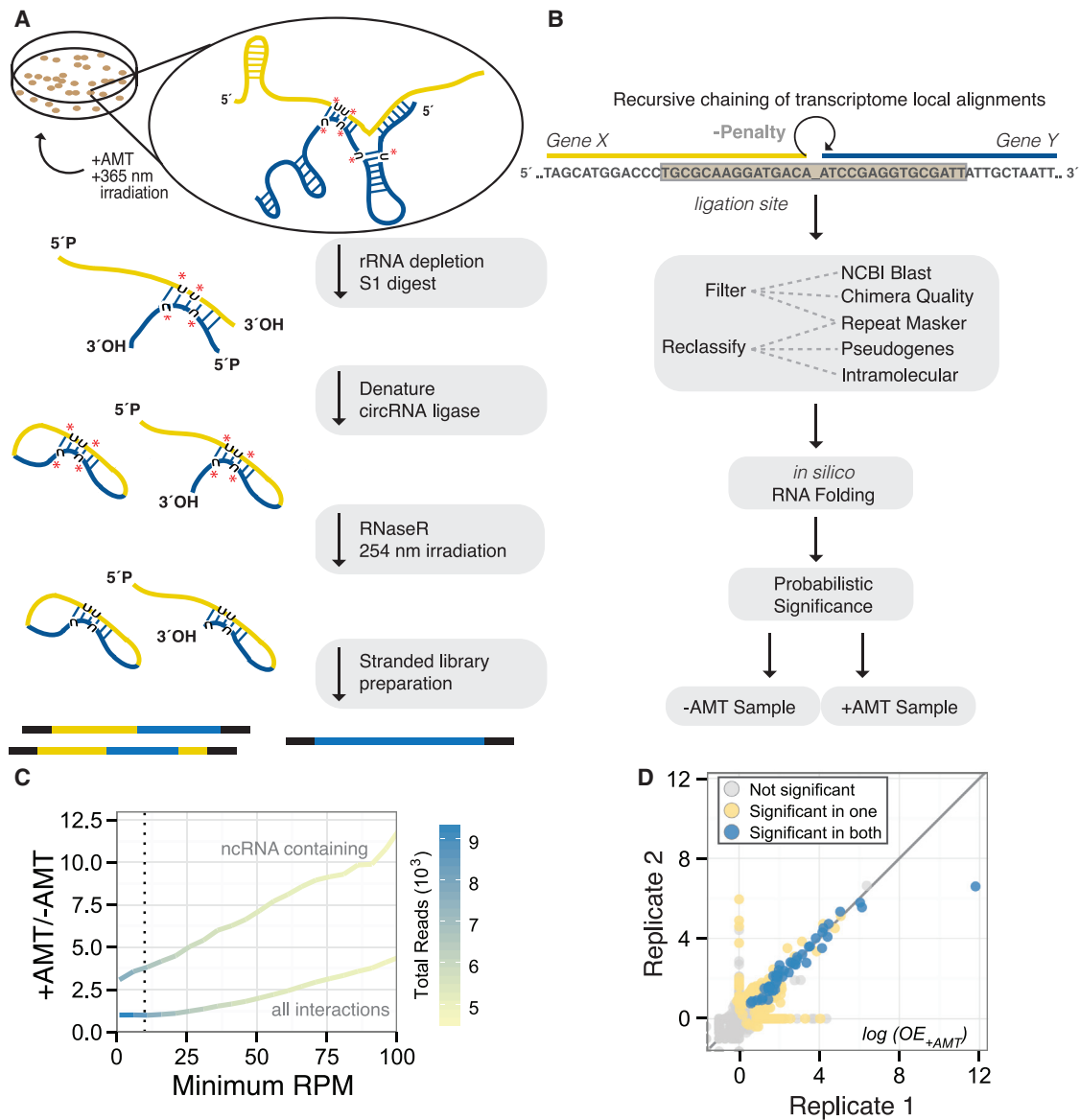
Figure 1. Overview of LIGR-Seq

(A) Schematic overview of the LIGR-seq protocol. Cells are incubated with AMT and irradiated with 365 nm UV light. RNA is extracted, DNase1 treated, and ribosomal RNA (rRNA) is depleted. Next, the RNA is digested with S1 nuclease-generating ends compatible for ligation by circRNA ligase. Crosslinked RNA is enriched using RNase R (Figures S1A and S1B) and crosslinks are reversed by 254 nm UV irradiation. Samples are prepared for strand-specific libraries for high-throughput sequencing. Control samples prepared in parallel are without the addition of AMT or ligase.

(B) Schematic overview of the "*Aligater*" analysis used to detect and score significance of detected RNA-RNA chimeric reads (Figures S1C–S1F and Experimental Procedures for details).

(C) Ratio of chimeric reads representing significant interactions in +AMT versus –AMT samples, in relation to a sliding cutoff for the minimum expression level of the chimera partner with the lowest expression. Interactions with observed/expected ratios that are higher in the +AMT sample compared to the –AMT sample are shown. RPM, reads per million total reads; color scale, number of significant chimeric reads in the +AMT sample ×10³.

(D) Correlation of the log fold ratio of observed versus expected values in the +AMT versus –AMT samples for significant (blue) and non-significant interactions (yellow) detected in both replicates. Interactions with observed/expected ratios (OE) that are higher in the +AMT sample compared to the –AMT sample are shown ($OE_{+AMT}/OE_{-AMT} > 1.1$; see Supplemental Experimental Procedures).

accessibility of duplexes or digested RNA free ends, may also affect the detection of true-positive interactions (see below). Nevertheless, in contrast to the specific detection of true-positive snRNA-snRNA interactions, only background levels of chimeras representing combinations of unexpected snRNA-snRNA interactions are detected (Figure 2A). Mapping of individual chimeric sequences onto previously defined secondary structures for U4+U6 snRNA and U2+U6 snRNA
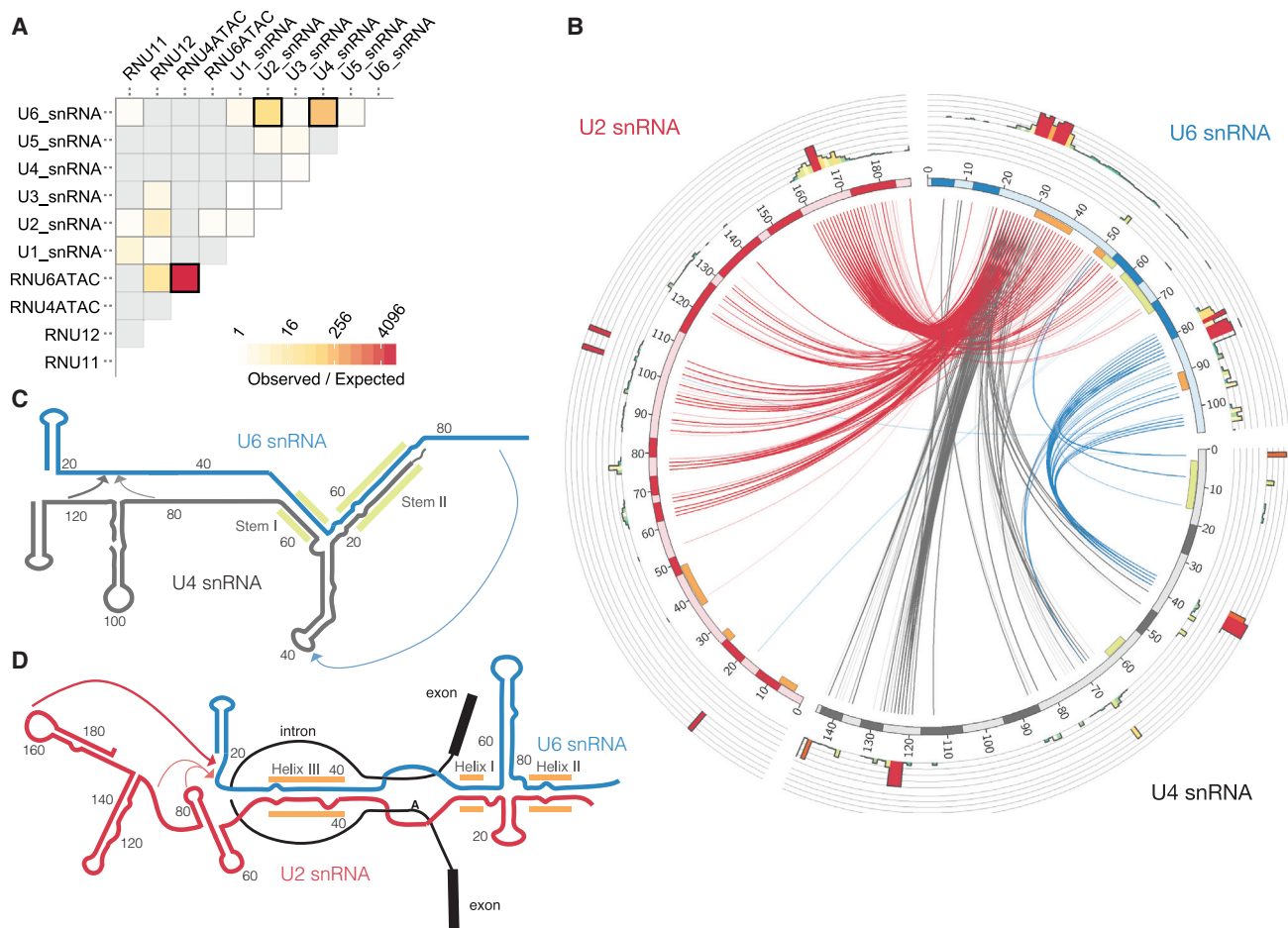
Figure 2. Specific Detection of In Vivo snRNA-snRNA Interactions by LIGR-Seq

(A) Spliceosomal snRNA-snRNA interactions detected using LIGR-seq. U3 snoRNA is analyzed as a negative (specificity) control. Color scale, $\log_2$ observed over expected ratio of chimeric read counts. Gray boxes indicate no supporting chimeric reads. Bolded boxes denote significant interactions.

(B) Circos plot showing intermolecular ligation junctions detected between U2+U6 snRNAs and U4+U6 snRNAs. Numbers indicate positions (in nt) within each snRNA. Red, blue, and gray ideograms indicate helical regions (dark shading) from known secondary structures involving U2, U4, and U6 snRNAs, respectively (Madhani and Guthrie, 1992; Rinke et al., 1985) and the inner circle green and orange colored boxes correspond to previously mapped (Madhani and Guthrie, 1992; Rinke et al., 1985) intermolecular interactions. Links within the circos plot indicate ligation sites between two snRNAs, with color corresponding to the snRNA upstream of the ligation site. The normalized histograms mark relative counts of chimera ligation sites, with the outer and inner histograms indicating snRNAs contributing a 5′ or 3′ end, respectively.

(C) Previously supported secondary structure model for the interaction between U4/U6 snRNAs, with major ligation junctions shown as links in (B) and mapped as colored arrows in (C).

(D) Same as in (C) but for the U2+U6 snRNA interaction.

interactions reveals non-uniform distributions of these ligation products that significantly coincide with expected single-stranded regions (Figure 2B; χ^2 goodness-of-fit test, p < 0.001). These data thus provide a map of accessible and proximal RNA sequences that form ligation products during the formation of spliceosomes (Figures 2C and 2D).

Although LIGR-seq employs an efficient rRNA depletion step (Figure 1; Experimental Procedures), many of the detected chimeric reads represent interactions associated with rRNA. This is due to the abundance (∼0.1% of total RNA) of remaining rRNA, but also because multiple steps of the LIGR-seq protocol, including AMT crosslinking and RNase R digestion, favor detection of highly structured RNAs such rRNAs and snRNAs. An

analysis of chimeras mapping to 28S, 18S, and 5S rRNAs revealed numerous RNA-RNA interactions with non-uniform distributions (Figures S2A and S2B) that reflect known secondary and tertiary interactions within the three-dimensional structure of the 80S ribosome (Figure S2C) (Anger et al., 2013). The most abundant +AMT-dependent chimeras representing intermolecular interactions involve 28S+5S rRNAs (OE_{AMT} = 6.4, p < 2.2e-16), and major ligation junctions cluster in known interaction regions (Figure S2A). In contrast, rRNA-rRNA interactions not known to occur in the ribosome (e.g., 5S+18S) were not detected above background levels. Taken together with the results from analyzing snRNA-snRNA interactions, these observations further demonstrate that LIGR-seq effectively captures

well-established examples of structural and functional RNA-RNA interactions with high specificity in vivo.

Global Snapshot of the RNA-RNA Interactome

We next systematically analyzed all LIGR-seq products to discover biologically important RNA-RNA interactions. Figures 3A and 3B display landscapes of RNA-RNA interactions that are significantly enriched beyond expected levels in the +AMT sample and, for comparison and control purposes, the –AMT sample, respectively. We detected an average of 514 significant interactions in the +AMT samples (Table S2). As expected, chimeras representing intermolecular rRNA-rRNA and snRNA-snRNA contacts were the most enriched compared to other classes of RNA-RNA interactions in the +AMT sample. We additionally detect numerous interactions between snoRNAs and other classes of ncRNAs, as well as with mRNA (Table S1; see below). In contrast, the most abundant classes of ligated product detected in the –AMT sample relative to the +AMT sample comprise mRNA+mRNA interactions, and less frequently mRNA+ncRNA interactions (Figure 3B; Table S1). The most abundant and stable interactions in vivo that are crosslinked by AMT thus involve at least one ncRNA, whereas mRNA-mRNA interactions are less frequently detected (Figures 1C, 2A, and 2B). These results further demonstrate that LIGR-seq data are highly enriched in specific, duplex-dependent interactions that form in vivo and thus indicate that it can be used to identify functionally important RNA-RNA interactions.

LIGR-seq reveals unexpected RNA-RNA interactions involving all functional classes of RNAs (Figure 3A; Table S1 for a full list). Here, we focus on further investigating snoRNA-RNA interactions. Aside from known snoRNA+snRNA and snoRNA+rRNA interactions, snoRNAs were observed to have the highest number of reads representing significant, previously unknown interactions in the +AMT sample. C/D box snoRNAs guide 2′-O-methylation of rRNA and snRNA via base-pairing interactions and are characterized by the presence of RUGAUGA (terminal box C) and CUGA (terminal box D and internal box D′) sequences (Figure 3C; Figure S3A) (Kiss, 2002). Sequences that form base-pairing interactions with target RNAs are located directly upstream of the D or D′ boxes (Matera et al., 2007). To investigate snoRNA+ncRNA and snoRNA+mRNA interactions detected by LIGR-seq, we predicted intermolecular base-pairing interactions proximal to junctions in chimeric reads (Figure S3B). In agreement with previous results (Kiss, 2002), known C/D box snoRNA interactions are associated with predicted duplexes that are >7 nt long and tend to be located immediately upstream of D or D′ boxes. In contrast, predicted snoRNA-snRNA duplexes among the set of previously unknown interactions are generally located at greater distances from D and D′ boxes (Figure 3C; p < 1.03e-6, Kolmogorov-Smirnov test). Many of these interactions involve orphan snoRNAs, as well as previously characterized snoRNAs engaged in predicted non-canonical interactions (Figure 3D). For example, LIGR-seq detects significant interactions between the orphan snoRNA SNORD89 and U2 snRNA, and between SNORD44 and a previously unknown target, the telomerase RNA component (TERC). Both ligations were validated by RT-PCR assays (Figures S3C and S3D). These results provide evidence that C/D box snoRNAs form interactions with a broader range of target RNAs than previously anticipated.

SNORD83B Controls the Levels of Target mRNAs

Orphan C/D box snoRNAs have previously been implicated in the regulation of mRNA stability, A-to-I editing, and alternative splicing through direct RNA-RNA interactions (Doe et al., 2009; Falaleeva et al., 2015, 2016; Kishore and Stamm, 2006). However, the range of possible functions of these ncRNAs is not known. To investigate the functional significance of previously unknown snoRNA-mRNA interactions, we focused on SNORD83B, an orphan box C/D snoRNA for which LIGR-seq detected multiple, +AMT-stabilized interactions with distinct mRNAs (Table S1 and Figure 4A). SNORD83B was efficiently knocked down using two independent modified antisense oligonucleotides (ASOs) (Figures 4B and S4A). Importantly, in both knockdowns, we detected significant increases in the steady-state levels of three of four SNORD83B target mRNAs, NOP14, RPS5, and SRSF3 (Figure 4C), but not of transcripts from the rpl3 gene, within which the SNORD83B locus resides. This indicates that the effect of the ASO knockdowns is due to SNORD83B depletion and not an indirect effect of RPL3 depletion (Figure S4B). Moreover, transcripts of comparable abundance from four genes that were not detected as SNORD83B targets also did not show significant changes in steady-state levels following knockdown with either ASO (Figure 4C). Finally, the ASO knockdowns also did not appear to affect transcriptional activity of the NOP14, RPS5, and SRSF3 genes, since the levels of unspliced transcripts from these genes were not appreciably altered (Figure 4D). Collectively, these results provide evidence that the orphan C/D box snoRNA SNORD83B controls steady-state levels of target mRNAs detected by LIGR-seq.

DISCUSSION

In this study, we describe LIGR-seq, a method that enables the generation of global-scale maps of RNA-RNA interactions in vivo. LIGR-seq complements recently described procedures for the transcriptome-wide mapping of RNA secondary structure in vivo (Ding et al., 2014; Rouskin et al., 2014; Spitale et al., 2015), long-range intramolecular RNA-RNA interactions in vitro (Ramani et al., 2015), and RNA interactions and structures associated with specific proteins of interest (Helwak et al., 2013; Kudla et al., 2011; Sugimoto et al., 2015). Similar to proximity ligation procedures developed to map long-range chromatin interactions (Dekker et al., 2002), LIGR-seq directly detects long-range interactions involving RNA molecules in cells. LIGR-seq thus reveals previously inaccessible information on RNA-RNA interactions that inform downstream functional and mechanistic studies.

Comparisons of significant chimeric reads detected by LIGR-seq with well-defined structural models of snRNA and rRNA intra- and intermolecular interactions demonstrate the specificity of the method and its ability to capture dynamic, functional interactions in vivo. In particular, the most significant LIGR-seq chimeric reads coincide with known single-stranded regions, where the ligation sites are generally located proximal to known
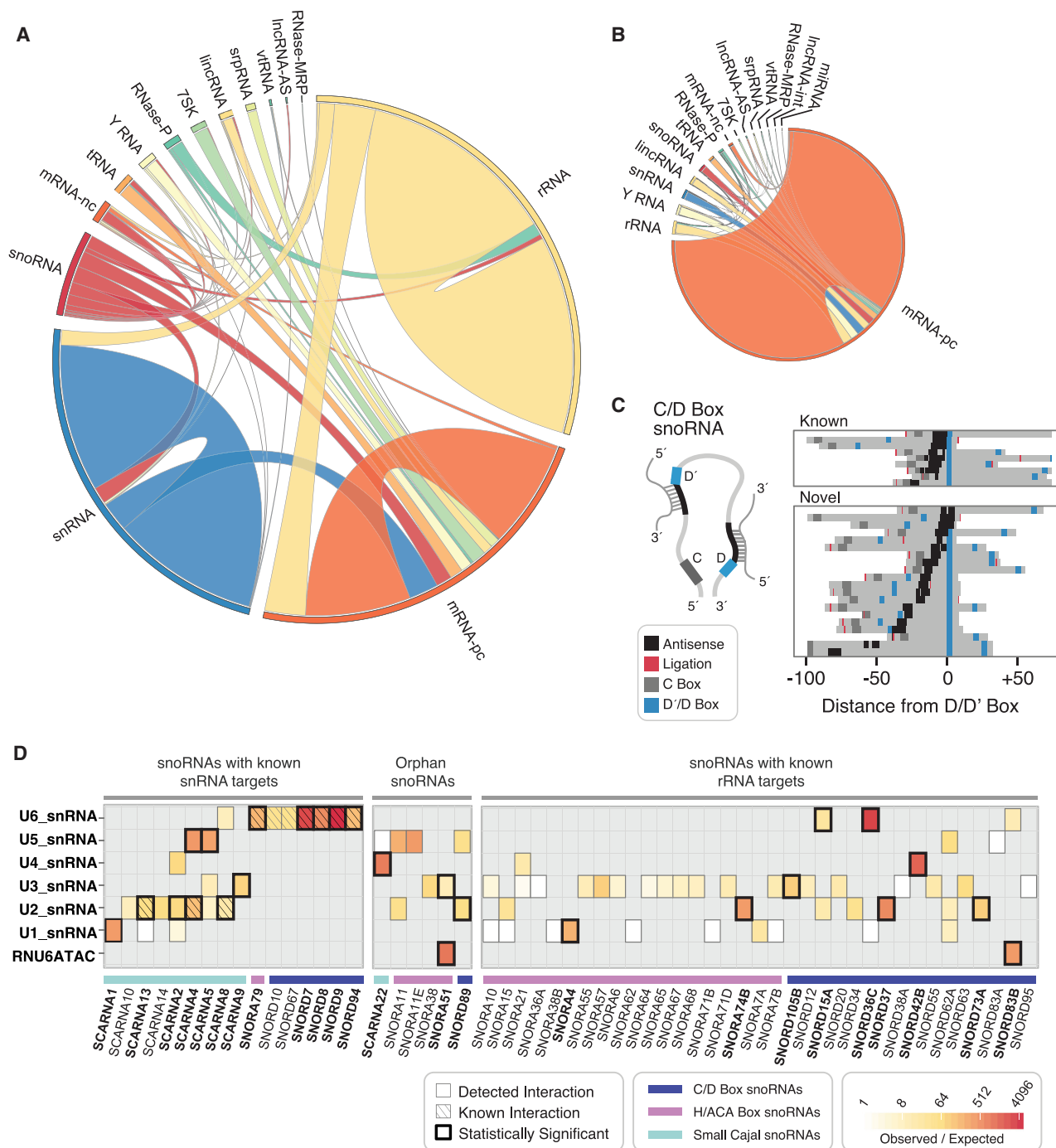
Figure 3. LIGR-Seq Analysis of the Human RNA-RNA Interactome

(A) Circos plot of the landscape of human RNA-RNA interactions detected by LIGR-seq. Link width for each class of RNA depicts the relative percent of chimeric reads representing significant interactions that have an observed versus expected ratio that is higher in the +AMT sample compared to the −AMT sample ($OE_{+AMT}/OE_{-AMT} > 1.1$). mRNA-pc, protein-coding mRNA; mRNA-nc, mRNAs lacking a predicted open reading frame; lncRNA-int, intronic lncRNAs.

(B) Same as in (A) but for interactions that have an observed versus expected ratio that is higher in the −AMT sample compared to the +AMT sample ($OE_{+AMT}/OE_{-AMT} < 0.9$).

(C) Predicted antisense regions (black) of snoRNAs on known and previously unknown targets detected by LIGR-seq. Each line represents one snoRNA and all snoRNAs have been aligned to a D/D′ box (blue). In red are positions of a ligation junction with the associated snRNA.

(D) Table of C/D box and H/ACA box snoRNA-snRNA interactions detected by LIGR-seq. Significant interactions are indicated with bold boxes. Hatching indicates a known interaction. Color scale, observed over expected ratio of chimeric reads.
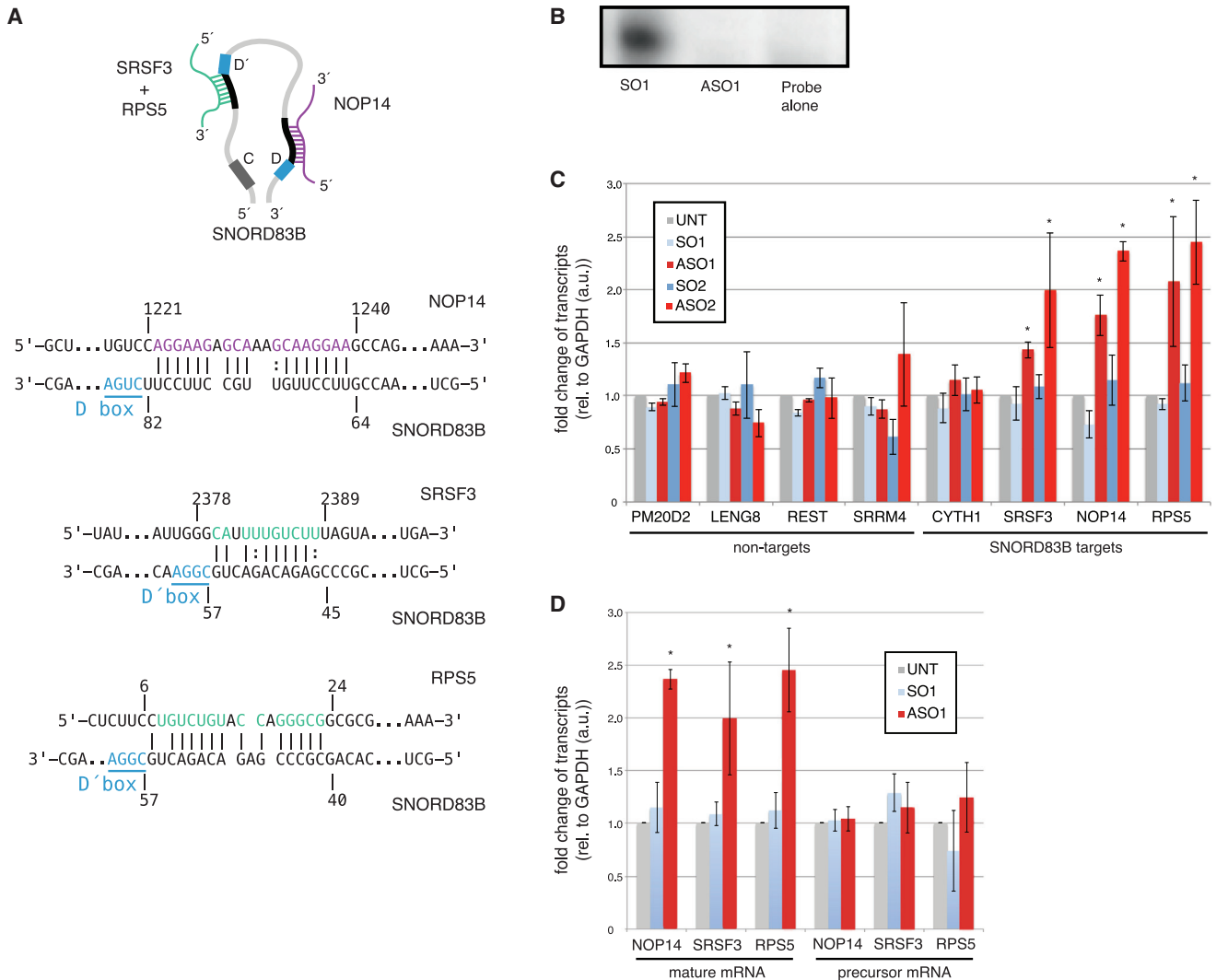
Figure 4. LIGR-Seq-Detected SNORD83B Interactions Affect mRNA Levels

(A) Predicted SNORD83B-target mRNA interactions. D/D′ boxes of the snoRNA are indicated in blue, and target mRNA sequences are indicated in green and purple.

(B) RNase protection assay (RPA) monitoring SNORD83B levels in 293T cells following transfection of modified oligonucleotides with sequences that are sense (SO) or antisense (ASO) to SNORD83B. "Probe alone," RPA probe following incubation with RNase but without input RNA.

(C) RT-qPCR assays monitoring levels of four mRNAs detected by LIGR-seq as SNORD83B-mRNA targets, and four control mRNAs of comparable abundance not detected as targets. Fold changes in mRNA level following ASO and control treatments are relative to Gapdh and normalized to untransfected controls (average of n = 3 biological replicates; error bars show SD; *p < 0.05, Welch's t test).

(D) RT-qPCR assay monitoring levels of mature and precursor SNORD83B transcripts. Fold changes in mRNA following ASO and control treatments are relative to Gapdh and normalized to untransfected controls (average of n = 3 biological replicates; error bars indicate SD; *p < 0.05, Welch's t test).

or predicted duplexes that most likely correspond to the sites of AMT crosslinking. The distances between the ligation junctions and these duplex regions are specific to each detected RNA-RNA interaction, and most likely are influenced by constraints presented by proximal secondary and tertiary structures that limit the extent of RNase digestion prior to the ligation step. In this regard, it should be noted that LIGR-seq does not efficiently detect RNA-RNA interactions involving short ncRNAs (e.g., miRNAs) since the corresponding duplexes do not yield sufficient lengths of free RNA ends to facilitate efficient ligation,

and short RNA sequences forming chimeras are difficult to unambiguously map by our computational method. While LIGR-seq does not directly map AMT-crosslinked duplexes, in most cases these can be inferred based on sequence complementarity proximal to ligation sites within significant chimeras. Moreover, the validity of these predicted sites of interaction is further strongly supported by the significant enrichment of the corresponding chimeras in +AMT versus control –AMT samples.

Among other interesting examples, LIGR-seq afforded the detection of many previously unknown interactions involving

snoRNAs and additional classes of RNAs in human cells. Previous work has linked snoRNAs to the regulation of alternative splicing, A-to-I editing, and transcript stability, although the precise mechanisms involved are unclear (Doe et al., 2009; Falaleeva et al., 2015; Kishore and Stamm, 2006; Falaleeva et al., 2016; Brameier et al., 2011). In the present study, we provide evidence that the orphan snoRNA SNORD83B controls the steady-state levels of several LIGR-seq-detected target mRNAs. Moreover, the detection of numerous additional snoRNA-mRNA chimeras by LIGR-seq suggests that there may be additional examples of snoRNA-mediated gene regulation. These results thus help guide future experiments directed at investigating specific mechanisms by which snoRNAs impact the regulation of mRNA transcripts.

In summary, our results demonstrate the utility of LIGR-seq in illuminating functionally important RNA-RNA interactions. Future application of LIGR-seq in mapping RNA-RNA interactions involving transcripts expressed over a wider range of abundance in diverse cell and tissue types, developmental stages, and in different species is expected to further facilitate efforts toward understanding the roles of the multitude of transcribed ncRNAs that currently lack known functions.

EXPERIMENTAL PROCEDURES

LIGR-Seq Protocol and Data Analysis
AMT Crosslinking and RNA Extraction
293T cells were cultured in DMEM with 10% FBS, aliquoted into 2×10^7 cells/tube and pelleted. Each pellet was resuspended in 2 mL cold TS buffer and AMT (Sigma) was added to a final concentration of 20 μg/mL. Cells were incubated for 10 min on ice, transferred to an open 60 mm petri dish placed on ice, and irradiated at 365 nm UV for 30 min at a distance of 15 cm from UV light source. RNA was extracted with TRI-reagent (Sigma) as per the manufacturer's instructions and treated with TURBO DNase I (Ambion). See Supplemental Experimental Procedures for full details.

RNA Preparation and Ligation
Ribosomal RNA was depleted from 4 μg of total RNA using the Ribozero Gold kit (Epicenter). 400 ng of RNA was digested for 30 min at room temperature in a 20 μL reaction in 1 × S1 buffer containing 2 μL S1 enzyme (diluted 1:100 in 1 × S1 buffer). Reactions were stopped by phenol-chloroform extraction. Ligation was performed using circRNA ligase (Epicenter) in 18 μL reactions containing 50 ng of S1-digested RNA and pre-incubated with 2 μL 10× circRNA ligase buffer for 2 min at 85°C to denature RNA. Tubes were immediately transferred to ice; 1 μL of 10 mM ATP and 1 μL of circRNA ligase was added and reactions were incubated in a thermocycler for 1 hr at 60°C.

Enrichment of Crosslinked RNA and Crosslink Reversal
Reactions were mixed with 0.5 μL of RNase R (Epicenter), 2.5 μL RNase R buffer, and 2 μL H_2O and incubated for 10 min at 37°C then stopped on ice. RNA was phenol-chloroform extracted, ethanol precipitated, washed in 70% ethanol, then resuspended in 20 μL of TS buffer and placed on ice in open Eppendorf tubes. RNA was irradiated for 10 min at 254 nm, ∼15 cm from the UV bulb. Samples were ethanol precipitated and prepared for sequencing (see Supplemental Experimental Procedures).

Data Analysis
Analysis of LIGR-seq data using *Aligater* involved local alignment and detection of chimeras, filtering and reclassification of chimeras, and a probabilistic analysis to define significant interactions (refer to Supplemental Experimental Procedures for details). Procedures for the analysis of C/D box snoRNA interactions are also described in Supplemental Experimental Procedures.

Functional Analysis of snoRNA-mRNA Interactions
SNORD83B was knocked down using modified ASOs (IDT) containing 2′-O-methyl and phosphorothioate-modified nucleotides. Details of knockdown conditions, RNAase protection and RT-qPCR assays used to characterize levels of SNORD83B and its target mRNAs are provided in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the sequence data reported in this paper is GEO: GSE80167.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2016.04.030.

AUTHOR CONTRIBUTIONS

E.S., T.S.-W., and B.J.B. designed the study, interpreted and analyzed data, and wrote the manuscript. D.O. assisted with experiments analyzing snoRNA function.

ACKNOWLEDGMENTS

REFERENCES

Anger, A.M., Armache, J.P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. Nature *497*, 80–85.

Behrens, S.E., and Lührmann, R. (1991). Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. Genes Dev. *5*, 1439–1452.

Belton, J.M., and Dekker, J. (2015). Chromosome Conformation Capture (3C) in Budding Yeast. Cold Spring Harb. Protoc. *2015*, 580–586.

Brameier, M., Herwig, A., Reinhardt, R., Walter, L., and Gruber, J. (2011). Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic Acids Res. *39*, 675–686.

Calvet, J.P., and Pederson, T. (1979). Heterogeneous nuclear RNA double-stranded regions probed in living HeLa cells by crosslinking with the psoralen derivative aminomethyltrioxsalen. Proc. Natl. Acad. Sci. USA *76*, 755–759.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306–1311.

Ding, Y., Tang, Y., Kwok, C.K., Zhang, Y., Bevilacqua, P.C., and Assmann, S.M. (2014). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature *505*, 696–700.

Doe, C.M., Relkovic, D., Garfield, A.S., Dalley, J.W., Theobald, D.E., Humby, T., Wilkinson, L.S., and Isles, A.R. (2009). Loss of the imprinted snoRNA mbii-52 leads to increased 5htr2c pre-RNA editing and altered 5HT2CR-mediated behaviour. Hum. Mol. Genet. *18*, 2140–2148.

Falaleeva, M., Surface, J., Shen, M., de la Grange, P., and Stamm, S. (2015). SNORD116 and SNORD115 change expression of multiple genes and modify each other's activity. Gene *572*, 266–273.

Falaleeva, M., Pages, A., Matuszek, Z., Hidmi, S., Agranat-Tamir, L., Korotkov, K., Nevo, Y., Eyras, E., Sperling, R., and Stamm, S. (2016). Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. Proc. Natl. Acad. Sci. USA *13*, E1625–E1634.

Feigon, J. (2015). Back to the future of RNA structure. RNA *21*, 611–612.

Grosswendt, S., Filipchyk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. Mol. Cell *54*, 1042–1054.

Hall, S.L., and Padgett, R.A. (1996). Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. Science *271*, 1716–1718.

He, L., and Hannon, G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. Nat. Rev. Genet. *5*, 522–531.

Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell *153*, 654–665.

Kehr, S., Bartschat, S., Stadler, P.F., and Tafer, H. (2011). PLEXY: efficient target prediction for box C/D snoRNAs. Bioinformatics *27*, 279–280.

Kishore, S., and Stamm, S. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science *311*, 230–232.

Kiss, T. (2002). Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell *109*, 145–148.

Kudla, G., Granneman, S., Hahn, D., Beggs, J.D., and Tollervey, D. (2011). Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc. Natl. Acad. Sci. USA *108*, 10010–10015.

Lui, L., and Lowe, T. (2013). Small nucleolar RNAs and RNA-guided post-transcriptional modification. Essays Biochem. *54*, 53–77.

Madhani, H.D., and Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. Cell *71*, 803–817.

Matera, A.G., Terns, R.M., and Terns, M.P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat. Rev. Mol. Cell Biol. *8*, 209–220.

Noller, H.F. (2006). Biochemical characterization of the ribosomal decoding site. Biochimie *88*, 935–941.

Oulas, A., Karathanasis, N., Louloupi, A., Pavlopoulos, G.A., Poirazi, P., Kalantidis, K., and Iliopoulos, I. (2015). Prediction of miRNA targets. Methods Mol. Biol. *1269*, 207–229.

Ramani, V., Qiu, R., and Shendure, J. (2015). High-throughput determination of RNA structure by proximity ligation. Nat. Biotechnol. *33*, 980–984.

Rinke, J., Appel, B., Digweed, M., and Lührmann, R. (1985). Localization of a base-paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking. J. Mol. Biol. *185*, 721–731.

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature *505*, 701–705.

Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T., and Chang, H.Y. (2015). Structural imprints in vivo decode RNA regulatory mechanisms. Nature *519*, 486–490.

Sugimoto, Y., Vigilante, A., Darbo, E., Zirra, A., Militti, C., D'Ambrogio, A., Luscombe, N.M., and Ule, J. (2015). hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. Nature *519*, 491–494.

Sun, J.S., and Manley, J.L. (1995). A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. Genes Dev. *9*, 843–854.

Tafer, H., Kehr, S., Hertel, J., Hofacker, I.L., and Stadler, P.F. (2010). RNAsnoop: efficient target prediction for H/ACA snoRNAs. Bioinformatics *26*, 610–616.

Tarn, W.Y., and Steitz, J.A. (1996). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science *273*, 1824–1832.

Underwood, J.G., Uzilov, A.V., Katzman, S., Onodera, C.S., Mainzer, J.E., Mathews, D.H., Lowe, T.M., Salama, S.R., and Haussler, D. (2010). FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nat. Methods *7*, 995–1001.

Vincent, H.A., and Deutscher, M.P. (2006). Substrate recognition and catalysis by the exoribonuclease RNase R. J. Biol. Chem. *281*, 29769–29775.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.

Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z., Zhang, J., Spitale, R.C., Snyder, M.P., Segal, E., and Chang, H.Y. (2014). Landscape and variation of RNA secondary structure across the human transcriptome. Nature *505*, 706–709.

Wassarman, D.A., and Steitz, J.A. (1993). A base-pairing interaction between U2 and U6 small nuclear RNAs occurs in > 150S complexes in HeLa cell extracts: implications for the spliceosome assembly pathway. Proc. Natl. Acad. Sci. USA *90*, 7139–7143.

Will, C.L., and Lührmann, R. (2005). Splicing of a rare class of introns by the U12-dependent spliceosome. Biol. Chem. *386*, 713–724.