Molecular Cell
Article

Long Noncoding RNAs with snoRNA Ends

Qing-Fei Yin,[1,4] Li Yang,[2,4] Yang Zhang,[1] Jian-Feng Xiang,[1] Yue-Wei Wu,[1] Gordon G. Carmichael,[3,*] and Ling-Ling Chen[1,*]
[1]State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology
[2]Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology
Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[3]Department of Genetics and Developmental Biology, University of Connecticut Stem Cell Institute, University of Connecticut Health Center, Farmington, CT 06030-6403, USA
[4]These authors contributed equally to this work.
*Correspondence: carmichael@nso2.uchc.edu (G.G.C.), linglingchen@sibcb.ac.cn (L.-L.C.)
http://dx.doi.org/10.1016/j.molcel.2012.07.033

SUMMARY

We describe the discovery of *sno-lncRNA*s, a class of nuclear-enriched intron-derived long noncoding RNAs (lncRNAs) that are processed on both ends by the snoRNA machinery. During exonucleolytic trimming, the sequences between the snoRNAs are not degraded, leading to the accumulation of lncRNAs flanked by snoRNA sequences but lacking 5′ caps and 3′ poly(A) tails. Such RNAs are widely expressed in cells and tissues and can be produced by either box C/D or box H/ACA snoRNAs. Importantly, the genomic region encoding one abundant class of *sno-lncRNA*s (15q11-q13) is specifically deleted in Prader-Willi Syndrome (PWS). The PWS region *sno-lncRNA*s do not colocalize with nucleoli or Cajal bodies, but rather accumulate near their sites of synthesis. These *sno-lncRNA*s associate strongly with Fox family splicing regulators and alter patterns of splicing. These results thus implicate a previously unannotated class of lncRNAs in the molecular pathogenesis of PWS.

INTRODUCTION

Small nucleolar RNAs (snoRNAs) are a family of conserved nuclear RNAs (about 70-200 nt) that are usually concentrated in Cajal bodies or nucleoli where they either function in the modification of snRNAs or rRNA, or participate in the processing of rRNA during ribosome subunit maturation (Boisvert et al., 2007; Kiss, 2001; Matera et al., 2007). There are several hundred known cellular snoRNAs and the great majority of these are encoded in the introns of protein-coding genes (Filipowicz and Pogacić, 2002). SnoRNAs are processed from excised and debranched introns by exonucleolytic trimming (Tycowski et al., 1993; Kiss and Filipowicz, 1995) and carry out their functions in complex with specific protein components, forming ribonucleoprotein complexes (snoRNPs) (Kiss, 2001). There are two main classes of snoRNAs: box C/D snoRNAs and box H/ACA snoRNAs, both of which serve as guide RNAs complementary to specific target sequences. Box C/D snoRNAs guide 2′-O-ribose methylation and box H/ACA snoRNAs guide pseudouridine modifications. Box C/D snoRNAs contain four conserved sequence elements: box C (RUGAUGA) and box D (CUGA) near the 5′ and 3′ termini, respectively, and an internal copy of each box, called C' and D'.

Some snoRNAs expressed from an imprinted region have been implicated in an important human disease, Prader-Willi Syndrome (PWS) (Sahoo et al., 2008; de Smith et al., 2009; Duker et al., 2010; Cassidy et al., 2012). While most autosomal genes are expressed biallelically, genomic imprinting causes some to be expressed only from either the paternal or maternal chromosome. The 15q11-q13 region is imprinted, leading to expression of the SNURF-SNRPN gene and downstream noncoding region from the paternal chromosome, while the UBE3A gene is expressed biallelically in most cells, but only from the maternal chromosome in neurons (Figure 1A). The long neuron-specific paternal transcript contains at least 148 exons (which along with their introns appear to have arisen by duplications followed by sequence divergence) and spans 470 kb (Runte et al., 2001). All paternal transcripts downstream of the SNRPN gene are noncoding and have been considered primarily as precursors for small RNAs (Cavaillé et al., 2000; Royo and Cavaillé, 2008; Runte et al., 2001). Since the minimal paternal deletion region associated with PWS (108 kb) removes only SNORD109A, the SNORD116 cluster of 29 similar snoRNAs and IPW, the most current published model is that SNORD116 deficiency is the primary cause of disease (Sahoo et al., 2008; de Smith et al., 2009; Duker et al., 2010). The function of SNORD116s is unknown. While most snoRNAs exhibit complementarity to rRNA or snRNA targets, it is noteworthy that even after numerous attempts, no noncoding RNA, pre-mRNA or mRNA targets for any of the SNORD116 snoRNAs from 15q11-q13 have been confirmed. These snoRNAs show minimal complementarity to rRNA or other RNAs so far and show greater homology to one another than to other snoRNAs. A few potential targets have been predicted bioinformatically, but none has been validated (Bazeley et al., 2008). The human and mouse SNORD116s are similar but not identical (even in putative targeting regions), but their deficiency in the mouse recapitulates some of the features of PWS (Skryabin et al., 2007). This has led to some confusion in the field as to what the precise molecular cause of PWS is.

Increasing lines of evidence have demonstrated that long noncoding RNAs (lncRNAs) are involved in gene regulation and human diseases (Wang and Chang, 2011; Chen and Carmichael,

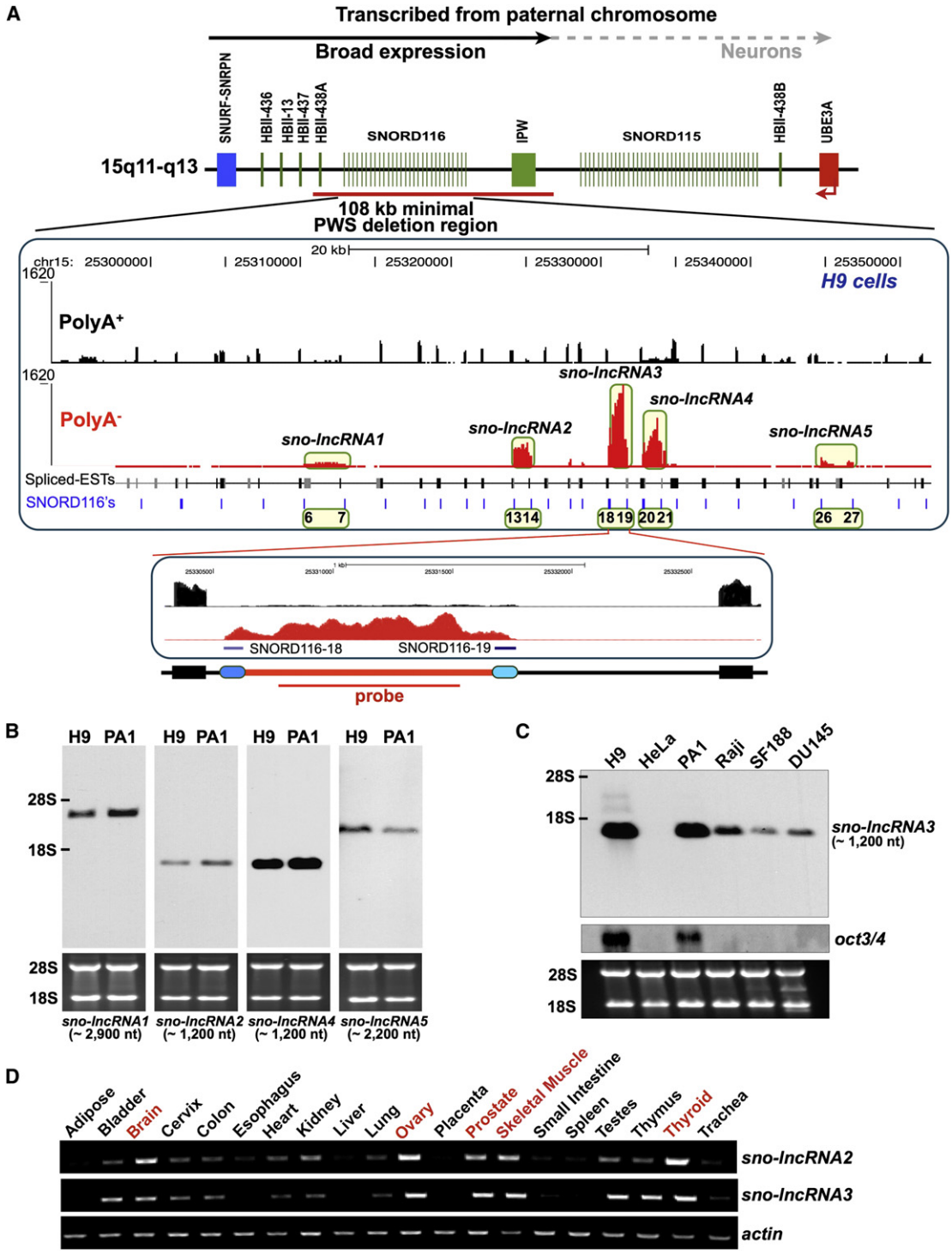

Figure 1. A Previously Unreported Class of Intron-Derived Long Noncoding RNAs from the Imprinted Region of chr15

(A) *sno-lncRNA*s from the imprinted region of chr15. Top, diagram of the region. Transcription of SNURF-SNRPN (blue) and downstream noncoding region (green) occurs only from the paternal chromosome and consists of one or two long primary transcripts containing more than 100 exons. UBE3A (red) is expressed bi-allelically in most cell types, but only from the maternal chromosome in neurons. The 15q11-q13 region is implicated in PWS and the minimal known deletion that results in the disease is shown. The PWS deletion region is known to express SNORD109A, the SNORD116 cluster of 29 snoRNAs and IPW. Middle, transcripts mapping to this region in H9 cells. Poly(A)$^+$ reads (black) map to known exons. Note the very high expression of long poly(A)- RNAs (red and boxed) that we have named *sno-lncRNA1-5*. The ends of these RNAs map to locations of the indicated snoRNAs. Bottom, a schematic enlarged view of *sno-lncRNA3*. Note that the ends of it map precisely to the intron-imbedded snoRNAs 18 and 19 from the SNORD116 cluster.

2010). The majority of lncRNAs are transcribed from intergenic regions (Guttman et al., 2009; Khalil et al., 2009), promoters (Hung et al., 2011) and enhancers (Ørom et al., 2010). Few, however, have been shown to derive from introns (for example, see Rearick et al., 2011; Salzman et al., 2012; Yang et al., 2011). Such molecules would be expected to lack both 5′ cap structures and 3′ poly(A) tails. We report here the discovery of a class of lncRNAs whose ends correspond to positions of intronic snoRNAs and which we have named snoRNA-related lncRNAs (*sno-lncRNA*s). We characterize the *sno-lncRNA*s from the PWS critical region of chr15 and demonstrate that at least some of these associate strongly with the splicing factor Fox2 and can alter splicing patterns in cells. Thus, these lncRNAs are implicated in the pathogenesis of an important human disease.

RESULTS

*Sno-lncRNA*s Are Produced from Introns with Two Imbedded snoRNA Genes

We recently used deep sequencing to investigate the nonpolyadenylated, or "poly(A)-," transcriptomes of HeLa cells and human embryonic stem (ES) cells (Yang et al., 2011). Specifically, we depleted both poly(A)$^+$ and ribosomal RNAs, performed size selection (>200 nt), and identified reads that uniquely aligned to the human genome. This allowed us to exclude from our analysis poly(A)$^+$ RNAs, rRNAs, abundant short RNAs (microRNAs, piRNAs, and siRNAs), tRNAs, small nuclear RNAs (snRNAs), many snoRNAs and repetitive transcripts such as the abundant *Alu* elements, LINE elements and endogenous LTRs. In this work we identified a number of intron-derived long noncoding RNAs that appeared to represent stable excised introns (Yang et al., 2011). Furthermore, after careful analysis we discovered an unusual class of lncRNAs (*sno-lncRNA*s) whose ends correspond to positions of intronic snoRNAs (Figures 1A and S1).

The most abundant *sno-lncRNA*s in human ES cells (H9) are expressed from the imprinted 15q11-q13 region (Figure 1A). Strikingly, in H9 cells at least five unannotated poly(A)- lncRNAs in the PWS region are expressed at extremely high levels (similar in abundance to some histone mRNAs) (Figure 1A). The deep sequencing data suggested that the ends of these molecules correspond precisely to intron-imbedded snoRNA genes and each molecule appears to derive from an intron containing two snoRNAs, one at each end (Figure 1A, bottom, and Figures S1–S3).

Northern blots (Figures 1B and 1C) demonstrated that full-length *sno-lncRNA* molecules of the expected sizes from this region are expressed in many different cell types, being especially abundant in pluripotent cells such as human ES H9 cells and ovarian carcinoma PA1 cells, but not expressed or expressed at low levels in HeLa cells (Figure 1C) and some other lines (Figure S4A). In addition, the PWS region *sno-lncRNA*s are widely expressed in many human tissues (Figure 1D), consistent with what is known of the expression from the 15q11-q13 region.

Since both classes of snoRNAs (box C/D and box H/ACA snoRNAs) are processed from excised and debranched introns by exonucleolytic trimming (Kiss and Filipowicz, 1995; Tycowski et al., 1993), we reasoned that introns containing two imbedded H/ACA box snoRNAs should also be capable of generating *sno-lncRNAs.* This is indeed the case (Figure S3), further highlighting the generality of this previously unreported class of intron-derived lncRNAs.

Box C/D *sno-lncRNA*s Are Processed Using the snoRNP Machinery

Deep sequencing and Northern blotting results suggested that the *sno-lncRNA*s from the 15q11-q13 region contain snoRNA sequences at each end. This is further supported by Northern blotting using probes to individual snoRNAs from this region. Figure 2A shows that probes specific for each of the snoRNAs flanking *sno-lncRNA2* not only reveal the annotated and reported snoRNAs (SNORD116-13 and SNORD116-14), but also the full-length *sno-lncRNA2*. We do not yet know whether this *sno-lncRNA* can serve as precursor for further processing to snoRNAs, or whether alternative splicing in this region (see Figure S1) leads to introns containing SNORD116-13 and SNORD116-14 individually.

Box C/D snoRNAs contain four conserved sequence elements: box C (RUGAUGA) and box D (CUGA) near the 5′ and 3′ termini, respectively, and a slightly less conserved internal copy of each box, called C' and D'. SnoRNAs are processed and carry out their functions in complex with specific protein components, forming ribonucleoprotein complexes (snoRNPs) (Kiss, 2001). Motif analysis of snoRNA sequences in all five PWS region *sno-lncRNA*s revealed they contain the conserved C/D boxes, but internal sequences between the snoRNAs are not very well conserved among the molecules (data not shown).

In conjunction with splicing, intron-encoded box C/D snoRNA sequences become associated with two of the four specific snoRNP proteins: 15.5K protein and Fibrillarin (Hirose et al., 2003). We therefore asked whether these proteins are associated with box C/D *sno-lncRNA*s. Immunoprecipitation confirmed that all five PWS region *sno-lncRNA*s can be immunoprecipitated with anti-Fibrillarin antibodies (Figure 2B), while a number of control RNAs, including a box H/ACA *sno-lncRNA,* are not. Furthermore, we expressed and purified a *sno-lncRNA* containing binding sites for the MS2 coat protein and found that both

(B) PWS region *sno-lncRNA*1, 2, 4 and 5 are highly expressed in human ES H9 and EC PA1 cells. Total RNA was isolated from each indicated cell line and resolved on an agarose gel. The probes for NB (red bar in A) were located in the internal sequence between two snoRNAs of each *sno-lncRNA*. Each *sno-lncRNA* from the PWS region can only be detected by an individual antisense probe and expression levels in these two cell lines are comparable. Equivalent amounts of RNA from H9 and PA1 cells were loaded as indicated by 28S and 18S rRNAs.
(C) The expression profile of *sno-lncRNA3* in different human cell lines by NB. *Oct3/4* was used as a maker for pluripotency (NB with the same stripped membrane is shown here). Note the *sno-lncRNA*3 expression is the highest in pluripotent cells.
(D) Widespread expression of PWS region *sno-lncRNA*s in human tissues by semiquantitative RT-PCR. The same amount of cDNA from each sample was used to amplify *actin* mRNA as a loading control. The red denoted tissues with a high expression level of *sno-lncRNA*s.

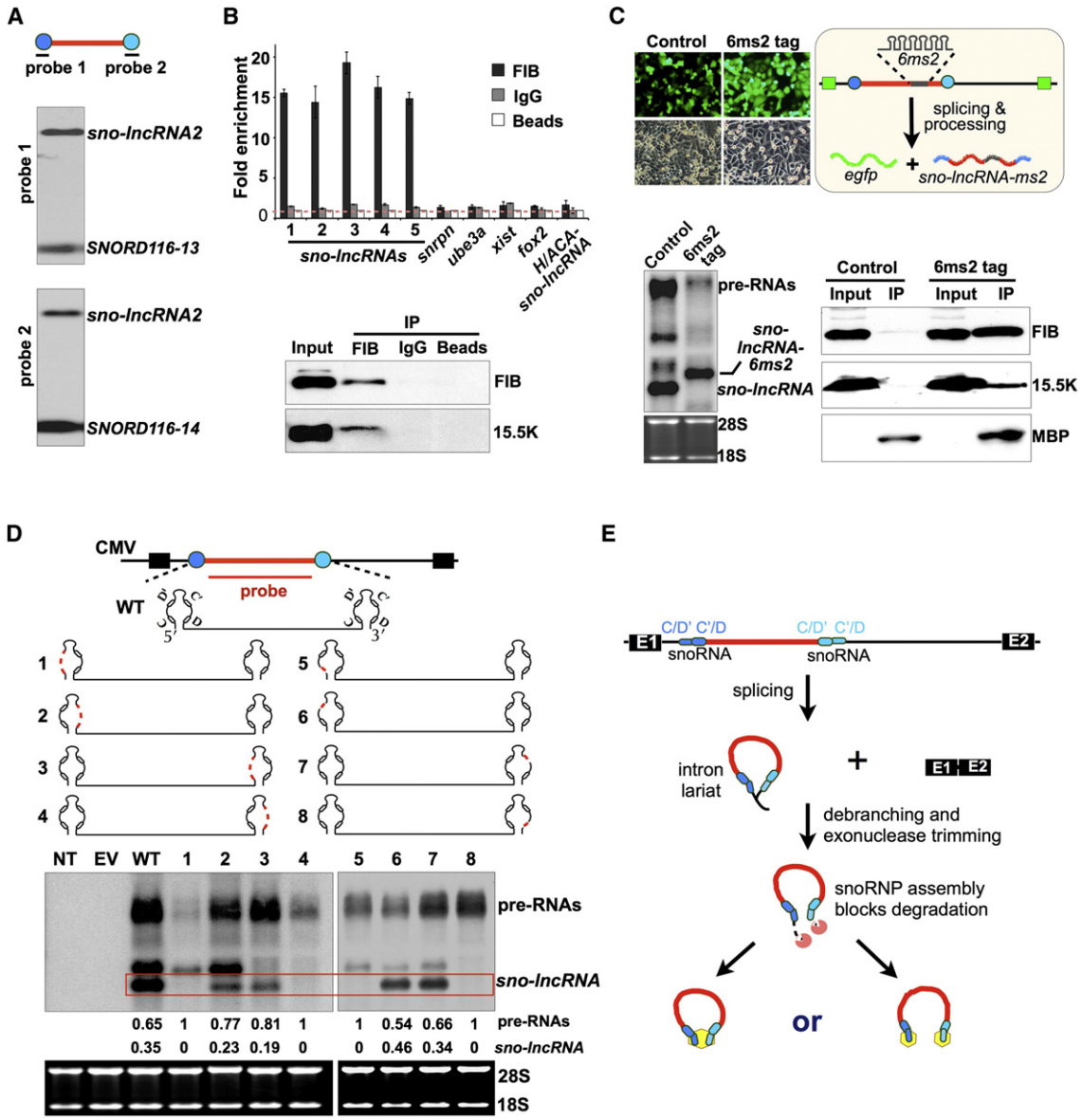

Figure 2. Processing of *sno-lncRNA*s

(A) snoRNAs and *sno-lncRNAs* can both be produced from the same genomic locus. The probes for NB are located in each snoRNA of *sno-lncRNA2* and these probes can detect both *sno-lncRNA2* and snoRNAs.

(B) PWS region *sno-lncRNA*s are associated with snoRNP complexes. Top, RIP was performed from PA1 cells using anti-Fibrillarin, IgG, and empty beads followed by RT-qPCR. Bar plots represent fold enrichments of RNAs immunoprecipitated by anti-Fibrillarin or IgG over empty beads, and error bars represent SD in triplicate experiments. Bottom, another key component of the snoRNP complex, 15.5K, associates with the same complexes.

(C) SnoRNP proteins are associated with *sno-lncRNA*s. Top left, *sno-lncRNA2* (for simplicity, referred to as *sno-lncRNA* in panels C and D) was inserted into the intron region of *egfp* mRNA such that only proper splicing leads to EGFP fluorescence (Top right, a schematic view of the vector used). Bottom left, total RNAs were collected from the same batch of transfected cells and resolved on agarose gels for NB. Note that both wt and the one containing an insert with 6ms2 were processed to stable *sno-lncRNA*s. Bottom right, MS2-MBP pulldown was performed from transfected HeLa cells, followed by WB with antibodies to Fibrillarin, 15.5K and MBP.

(D) *Sno-lncRNA* processing depends on C/D boxes of snoRNAs. Top, a schematic drawing of a wt *sno-lncRNA* in the expression vector. A full-length *sno-lncRNA* flanked by its natural intron, splicing sites and exons was cloned downstream of a CMV promoter. C/D' and C'/D boxes of each snoRNA and the 5′/3′ ends of *sno-lncRNA* are indicated; the red dashed lines represent deletions (clones 1 to 8). Bottom, HeLa cells do not express the PWS region *sno-lncRNA*s and hence were used for studies of mutant *sno-lncRNA* processing. Total RNA isolated from HeLa cells transfected with each indicated plasmid was resolved on agarose gels for NB with an antisense probe located in the internal sequence (red bar), and rRNAs were used as the loading control. NT, no transfection; EV, transfection with empty vector; WT, transfection with the wt *sno-lncRNA* in the vector; 1, deletion of C and D' boxes of snoRNA in the 5′ end; 2, deletion of C' and D boxes of snoRNA in the 5′ end; 3, deletion of C and D' boxes of snoRNA in the 3′ end; 4, deletion of C' and D boxes of snoRNA in the 3′ end; 5, deletion of C box only of snoRNA in the 5′ end; 6, deletion of D' box only of snoRNA in the 5′ end; 7, deletion of C' box only of snoRNA in the 3′ end; 8, deletion of D box only of

Fibrillarin and 15.5K protein can be copurified with the *sno-lncRNA* (Figure 2C). Taken together, these data suggest that *sno-lncRNA*s are generated using the same machinery involved in snoRNP biogenesis and that snoRNP components remain associated with them.

In order to learn more about the mechanism of processing of *sno-lncRNA*s, we constructed *sno-lncRNA* expression vectors in which the C and D' boxes or the C' and D boxes of the snoRNAs at both ends were deleted (Figures 2D). Results clearly show that the C and D' boxes of the snoRNA at the 5′ end and the C' and D boxes of the snoRNA at the 3′ end are critical for both *sno-lncRNA* processing and pre-RNA stabilization (Figure 2D, lanes 1 and 4), while other deletions still result in the proper processing of *sno-lncRNA*s, although with reduced efficiency (Figure 2D, lanes 2 and 3). Further mutagenesis revealed that the deletion of the C box at the very end of the 5′ snoRNA and the D box at the very end of the 3′ snoRNA completely block *sno-lncRNA* generation, while pre-RNA expression remains unchanged (Figure 2D, lanes 5 and 8). These data further confirm that the processing of *sno-lncRNA*s requires essential motifs of snoRNA molecules at each end.

Analysis of histone modifications of the SNRPN region using the ENCODE ChIP-Seq from human ES cells (H1) and neuronal NH-A cells revealed that *sno-lncRNA*s likely do not contain their own promoters but rather derive from a longer *SNRPN* transcript (Figure S4C). Taken together with the known processing of snoRNAs from excised, debranched introns by exonuclease trimming, these considerations lead us to propose the model for *sno-lncRNA* processing shown in Figure 2E. In this model, introns that contain two snoRNAs are processed from their ends; however, the internal sequences between snoRNAs are not removed, leading to the accumulation of long noncoding RNAs with snoRNP ends. Processing could occur in two ways, leading to linear RNAs with snoRNP structures at both ends, or "circularized" molecules resembling normal snoRNP structures but with very long internal inserts (Figure 2E). While our results are also consistent with the "circularized" model, we favor the model where naturally occurring *sno-lncRNA*s have canonical snoRNP structures at both ends. The presence of snoRNPs at each end could enhance the stability of these RNAs by inhibiting exonucleolytic degradation. In fact, we found these PWS region *sno-lncRNA*s to be quite stable. After treatment of PA1 cells with alpha-amanitin or Actinomycin D, we observed *sno-lncRNA*s 2,3 and 4 to remain virtually unchanged for at least 24 hr. *Sno-lncRNA*s 1 and 5 were somewhat less stable, with half lives of 12-16 hr (data not shown), consistent with their lower relative abundances in PA1 cells.

Localization of PWS Region *sno-lncRNA*s

Since they are processed from introns and are not polyadenylated (Figures 1, S1, S3, and S4B), we expected *sno-lncRNA*s to be localized to the nucleus. By nuclear/cytoplasmic RNA fractionation (Figure S4D) and RNA in situ hybridization, we found that the PWS region *sno-lncRNA*s are nuclear retained and each exhibits striking localization to primarily a single or two closely positioned subnuclear sites in multiple cell lines (Figures 3, S5A, and S5B). Intriguingly, our results also revealed that all five PWS region *sno-lncRNA*s accumulate to the same places in the nucleus (Figures 3A and S5A). This is consistent with a common function and/or a common fate of these *sno-lncRNA*s. Most snoRNPs localize to the nucleolus, which may reflect their normal targeting of rRNAs (Boisvert et al., 2007; Kiss, 2001; Matera et al., 2007), and many C/D box snoRNAs, proteins and Fibrillarin are also enriched in Cajal bodies (CBs). However, *sno-lncRNA*s do not accumulate in either nucleoli or CBs (Figure 3B), suggesting that they are functionally distinct from classical C/D box snoRNAs. This suggests that these *sno-lncRNA*s do not traffic through or encounter CBs or nucleoli as is the case for canonical snoRNAs. Double DNA/RNA FISH (Figure 3C) suggests that PWS region *sno-lncRNAs* accumulate at or near their sites of processing. In both H9 and PA1 cells, *sno-lncRNA*s colocalize with the site of one of their genes, consistent with the fact that they are expressed only from the paternal chromosome. Interestingly, when expressed from a plasmid expression vector, *sno-lncRNA* localizes to numerous nuclear sites, likely representing sites of plasmid transcription (Figure S5C).

Knockdown of PWS Region *sno-lncRNA*s Has Little Effect on Global Gene Expression

As *sno-lncRNA*s represent an unannotated class of molecules, we were interested in exploring their possible functions. Their exceptionally high expression in some cells and their localization to sites of their processing first suggested a possible role in chromatin organization or local gene expression. Since all PWS region *sno-lncRNA*s accumulate together, we optimized different combinations of phosphorothioate-modified antisense oligodeoxynucleotides (ASOs) (Figures 4 and S6) to knock down all five *sno-lncRNA*s simultaneously. We were able to knockdown all five to about 20%–30% of their normal levels in PA1 cells (Figure 4A). Figure 4B further illustrates the knockdown of *sno-lncRNA2*. This significant but incomplete knockdown resulted in fewer and smaller nuclear dots seen by FISH: about 60% of cells in which *sno-lncRNA*s were knocked down completely lost detectable nuclear accumulations (Figure 4C), while 40% of these cells contained smaller nuclear dots compared to control cells (Figures 4D and S6D). Further, knocking down *sno-lncRNA*s had no significant effect of the formation of snoRNAs (Figure S6E). If these lncRNAs played a role in gene expression from the PWS region, we would have expected to observe an altered level of *snurf-snrpn* transcription. However, this was not the case and we observed only slight decrease of expression across the SNURF-SNRPN gene (Figure 4E). This suggests that although they primarily accumulate to their sites of synthesis (Figure 3C), the PWS region *sno-lncRNA*s could be functionally distinct from other lncRNAs, such as *xist,* which

snoRNA in the 3′ end. Note that deletions of #1, 4, 5, and 8 impede *sno-lncRNA* processing (red box). The upper band of *sno-lncRNA* in NB is from aberrant splicing (data not shown). The relative abundance of pre-RNA and processed *sno-lncRNA* from each transfection was determined using Image J and labeled underneath.

(E) A model for *sno-lncRNA* processing. See text for details.

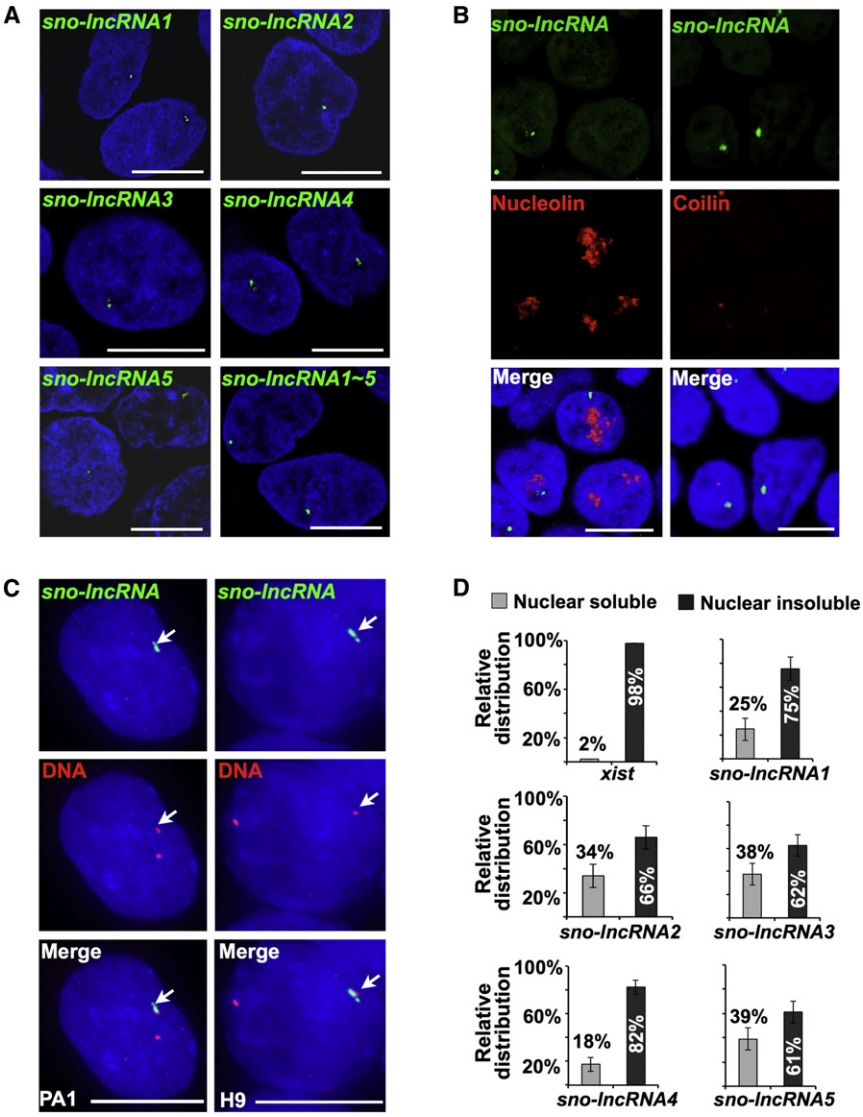

Figure 3. Nuclear Localization of *sno-lncRNAs*

(A) Each PWS region *sno-lncRNA* exhibits strong nuclear accumulation in PA1 cells. RNA ISH (green) was performed with individual probes recognizing the internal sequence of each *sno-lncRNA,* or with pooled probes recognizing all PWS region *sno-lncRNA*s (bottom right). Representative images are shown. DAPI is in blue and the white scale bar in all images denotes 10 μm.

(B) *Sno-lncRNA*s do not accumulate in either nucleoli or CBs. Costaining of *sno-lncRNA*s (green) and marker proteins for nucleoli (Nucleolin, red) and CBs (Coilin, red), respectively in H9 cells. Representative images are shown.

(C) *Sno-lncRNA*s accumulate at a single chromosomal locus. Double FISH of *sno-lncRNA*s (green) and its adjacent DNA region (red) in both PA1 and H9 cells. A single Z stack of representative images acquired with an Olympus IX70 DeltaVision Deconvolution System microscope is shown.

(D) Total RNA from PA1 cells was separated into cytoplasmic (not shown), nuclear soluble, and nuclear insoluble fractions. Bar plots represent relative abundance of RNAs in the nuclear soluble and insoluble fractions as measured by RT-qPCR. Error bars represent SD in triplicate experiments.

have a profound role in regulating gene expression *in cis* by recruiting chromatin modeling complexes (Augui et al., 2011). In fact, these PWS region *sno-lncRNA*s are less tightly associated with chromatin when compared to *xist,* as revealed by subnuclear fractionation (Figure 3D), further supporting the notion that these *sno-lncRNA*s exert their functions differently from many other known chromatin associated lncRNAs.

To further examine whether these *sno-lncRNA*s could affect gene expression from other regions of the genome, RNAs from duplicate experiments using scrambled or specific ASO-treated cells were collected for RNA-seq analyses. Since the correlation between two biological repeats was greater than $r^2 = 0.95$ (Figure S7A), we combined data from replicates for further analysis. Interestingly, however, in this entire set of data, we identified only a very small number of genes whose expression changed significantly after *sno-lncRNA* knockdown (Figure 4F and Tables S1 and S2). Importantly, these changes observed in the RNA-seq data were further validated by RT-qPCR (Figure 4G). Interestingly, one of the most affected genes, BIRC6, is known to play a crucial function in cell division (Hao et al., 2004) and promotes hippocampal neuron survival (Sokka et al., 2005), suggesting a possible role in PWS neurocognitive defects.

PWS Region *sno-lncRNA*s Are Associated with Splicing Factor Fox2

Further insights into the potential function of *sno-lncRNA*s was suggested by cross-linking and immunoprecipitation followed by deep sequencing (CLIP-seq) data from human ES cells using antibodies to Fox2, which revealed that transcripts from the *SNORD116* region are strongly bound by Fox2 (Yeo et al., 2009). The Fox family of alternative splicing regulators contains three paralogs: Fox1 (also called A2BP1), Fox2 (also called RBM9) and Fox3 (also called HRNBP3 and NeuN). Fox2 is widely expressed in tissues, but Fox1 and Fox3 are most highly expressed in the brain (Kim et al., 2011; Nakahata and Kawamoto, 2005; Underwood et al., 2005). These proteins act as alternative splicing regulators and have been shown to act by binding to the hexanucleotide motif UGCAUG (Kim et al., 2009; Underwood et al., 2005), but can also recognize GCAUGU, GUGAUG, UGGUGA and GGUGGU (Yeo et al., 2009). In addition, these factors may be involved in alternative polyadenylation, mRNA stability, localization and translation (Shi et al., 2009; Wang et al., 2008). Interestingly, each *sno-lncRNA* from the 15q11-q13 region contains multiple binding sites for Fox2 and one of the very strongest regions of

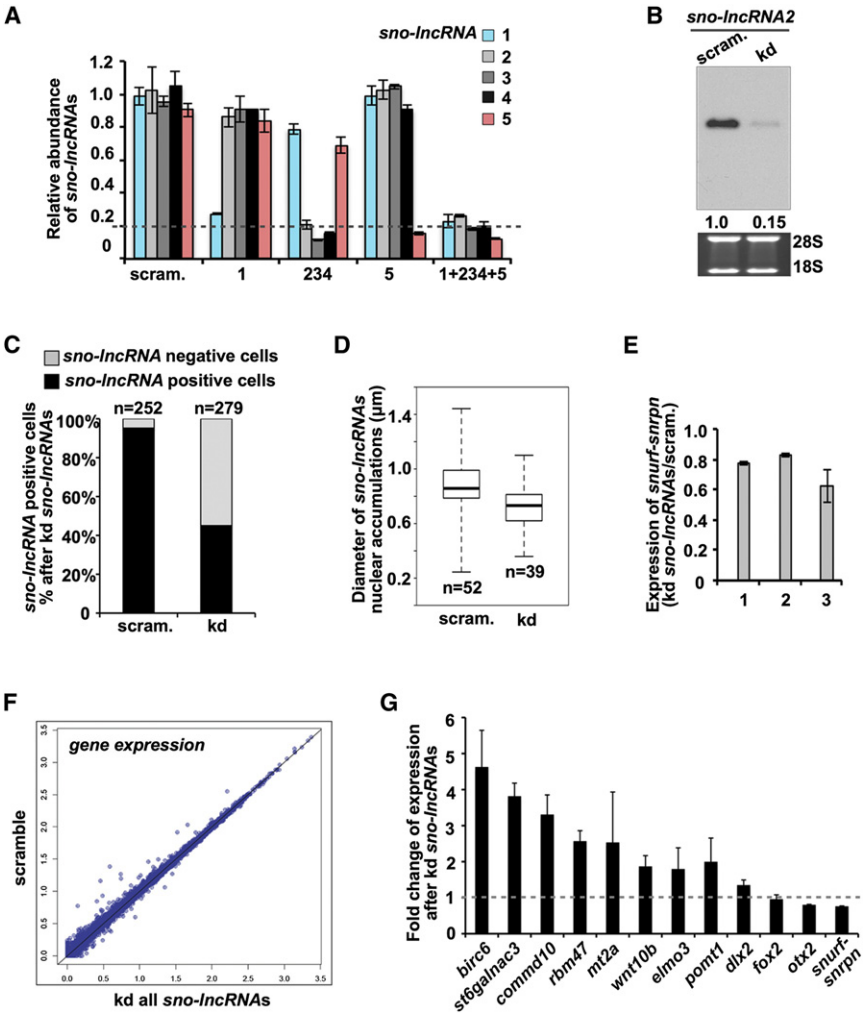

Figure 4. The Effect of Knockdown of *sno-lncRNA*s on Global Gene Expression

(A) Knockdown of *sno-lncRNA*s in PA1 cells. RT-qPCR indicates efficient knockdown of *sno-lncRNA* levels with single or pooled antisense ASOs (see Figures S6A–S6C for details).

(B) NB revealed efficient knockdown of *sno-lncRNA*s in PA1 cells. The percentage of *sno-lncRNA2* in the ASO treated cells is indicated underneath.

(C) Knockdown of all *sno-lncRNA*s leads to the disappearance of *sno-lncRNA*s nuclear accumulations in most cells. The diameters of *sno-lncRNA* nuclear accumulations in the remaining cells were calculated in (D).

(D) Boxplots showing the distribution of the diameters of *sno-lncRNA* nuclear accumulations in scramble ASO or pooled ASOs-treated cells.

(E) Knockdown of all *sno-lncRNA*s has no dramatic effect on *snurf-snrpn* expression as evaluated by different sets of primer pairs across SNURF-SNRPN by RT-qPCR.

(F) The effect of knockdown of all *sno-lncRNAs* on global gene expression. Correlation between concatenated biological repeats of scramble ASO and pooled ASOs treated PA1 cells for gene expression. The scale for all plots is log10 (See Tables S1 and S2 for details.)

(G) Knockdown of all *sno-lncRNA*s altered the expression of several genes by RT-qPCR.

In (A), (E) and (G), bar plots represent relative expression of genes (normalized to *actin*), and error bars represent SD in triplicate experiments.

interaction of Fox2 in the entire genome (Yeo et al., 2009) mapped precisely to *sno-lncRNA1* in H9 cells (Figures 5A, S7B). We confirmed that Fox2 directly interacts with these *sno-lncRNA*s using immunoprecipitation of Fox2 followed by RT-qPCR (Figure 5B) and by purification of a *sno-lncRNA* containing binding sites for the MS2 coat protein followed by Western blotting with Fox2 antibodies (Figure 5C). Additional colocalization of *sno-lncRNA* and Fox2 in both H9 cells and PA1 cells revealed that the nuclear Fox2 is strongly enriched in nuclear accumulations containing *sno-lncRNA*s (Figure 5D), suggesting binding of *sno-lncRNA*s could locally redistribute Fox2 to specific subnuclear neighborhoods.

PWS Region *sno-lncRNA*s Alter Fox-Regulated Splicing

The strong association of Fox2 with *sno-lncRNA*s suggested at least two possibilities. First, Fox2 may be necessary for the expression of at least some *sno-lncRNAs*. In this model, the splicing events that give rise to *sno-lncRNA*s may themselves rely on, or be influenced by, the expression of Fox proteins. Thus, these splicing factors could alter splicing patterns to make different *sno-lncRNA* expression. Fox2 but not Fox 1 or Fox3 is expressed in H9 (Yeo et al., 2009) and PA1 cells (Figure 5E and Table S1); however, knockdown of Fox2 (Figure 5F) had no significant effect on *sno-lncRNA* expression (Figure 5G).

The second formal possibility is that *sno-lncRNA*s influence Fox2-mediated splicing regulation. Given the facts that each *sno-lncRNA* contains multiple consensus binding sites for Fox2 (Figures 5A and S2), that the expression level of these *sno-lncRNA*s is high in pluripotent cells and in neurons (Figure 1), and that all five PWS region *sno-lncRNA*s accumulate to the same places in the nucleus (Figures 3, 5D, and S5), we hypothesized that at least some *sno-lncRNA*s may act as molecular sinks that titrate Fox proteins in specific cells thereby inducing subtle alterations in splicing patterns of Fox-regulated exons. To test the hypothesis that PWS region *sno-lncRNA*s may functionally alter Fox-regulated splicing, we used our RNA-seq data to analyze the splicing patterns of Fox-regulated cassette exons as reported previously (Zhang et al., 2008) and found that after knockdown of PWS region *sno-lncRNA*s some of these cassette exons exhibited consistent changes, while the level of expression of each gene remained unaltered (Figures 6A and S7C). More importantly, a number of altered exons were affected in opposite ways by depletion of PWS region *sno-lncRNA*s or by knockdown of Fox2 (Figure 6A), further supporting the hypothesis that PWS region *sno-lncRNA*s act as molecular sinks

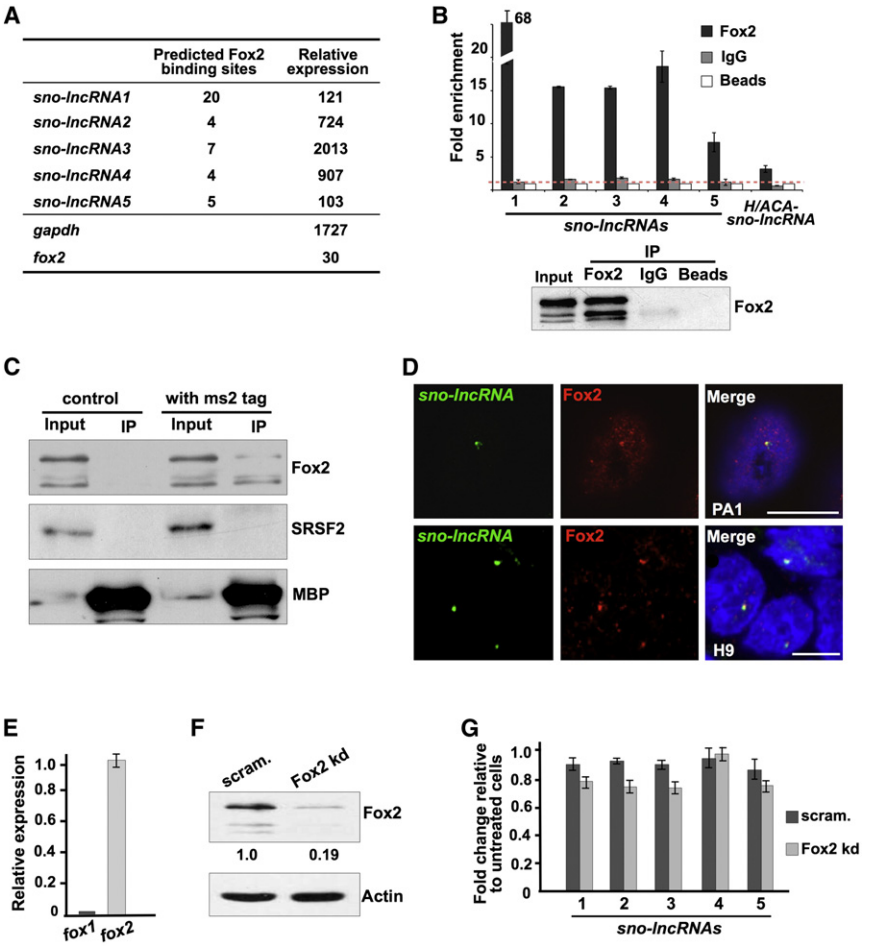

Figure 5. PWS Region *sno-lncRNA*s Interact with Splicing Factor Fox2

(A) *Sno-lncRNA*s contain multiple consensus binding sites for Fox family splicing factors (Yeo et al., 2009). The relative expression of each *sno-lncRNA* in H9 cells is shown (normalized value from the Solexa deep sequencing; Yang et al., 2011). The relative expression of *gapdh* and *fox2* in the same sample is also shown.

(B) PWS region *sno-lncRNA*s are associated with Fox2. RIP was performed from PA1 cells using anti-Fox2 antibody, IgG, and empty beads, followed by RT-qPCR analyses. Bar plots represent fold enrichments of RNAs immunoprecipitated by anti-Fox2 or IgG over empty beads, and error bars represent SD in triplicate experiments.

(C) Fox2 is associated with PWS region *sno-lncRNA*s. MS2-MBP pull-down was carried out as described in Figure 2C. *Sno-lncRNA-6ms2* specifically brings down Fox2, but not the general splicing factor SRSF2.

(D) *Sno-lncRNA*s localize to the enriched Fox2 accumulations in the nucleus. Colocalization of *sno-lncRNA*s (green) and Fox2 (red) in both PA1 and H9. Representative images are shown.

(E) PA1 cells express Fox2 but not Fox1 as examined by RT-qPCR.

(F) Knockdown of Fox2 is achieved using shRNA against Fox2 as confirmed by WB. The percentage of Fox2 in the shRNA treated PA1 cells is indicated underneath.

(G) *Sno-lncRNA* expression does not depend on Fox2. The relative abundance of *sno-lncRNA*s was analyzed by RT-qPCR.

In (E) and (G), bar plots represent relative expression of genes (normalized to *actin*), and error bars represent SD in triplicate experiments.

for Fox2. The fact that we have so far observed only subtle splicing alterations is quite likely due both to an incomplete simultaneous knockdown of all five *sno-lncRNA*s (Figure 4A) and to the fact that *sno-lncRNA* expression relocalizes some but not all nuclear Fox2 (Figure 5D). Taken together, the current results are consistent with a model in which the binding of *sno-lncRNA*s to Fox2 act locally to alter splicing patterns in specific subnuclear neighborhoods rather than throughout the nucleus (Figure 6B).

DISCUSSION

We report the discovery of a previously unannotated class of nuclear RNA molecules which are processed using the endogenous snoRNA biogenesis machinery. These *sno-lncRNA*s most likely eluded previous detection because they are long, derive from introns and lack poly(A) tails. Some of the *sno-lncRNA*s map to a genomic location responsible for Prader-Willi Syndrome, raising the possibility that understanding their function may lead to distinct insights into the pathogenesis of this disease. The most abundant *sno-lncRNA*s in pluripotent cells are those from the PWS region and are processed from five introns, each of which contains two snoRNA genes. Interestingly, while it has been reported that efficient box C/D type snoRNP processing often requires that the snoRNA gene be positioned near the 3′ splice site (Hirose et al., 2003; Hirose and Steitz, 2001), this is clearly not the case for any of the PWS region snoRNAs. In each case, the proximal snoRNA gene is positioned close to the 5′ splice site. In contrast, the distal snoRNA genes are far from the 3′ splice sites. We do not yet know whether these characteristics are mechanistically connected to the efficiency of processing or stability of the PWS region *sno-lncRNA*s.

While snoRNAs are known to localize to Cajal bodies or nucleoli, this is not true for the PWS region *sno-lncRNA*s (Figure 3B). The colocalization revealed by DNA and RNA double FISH (Figure 3C) demonstrates an accumulation of these molecules near their sites of synthesis. However, as these *sno-lncRNA*s are less tightly associated with chromatin (Figure 3D), their localization could also reflect previously unknown subnuclear structures that remain to be dissected. Also, it is quite possible that *sno-lncRNA*s expressed from sites elsewhere in the genome may not only have different molecular functions, but may localize to different sites within the nucleus. In addition, the nonpolyadenylated RNA sequencing from HeLa and H9 cells revealed a differential expression pattern of the PWS

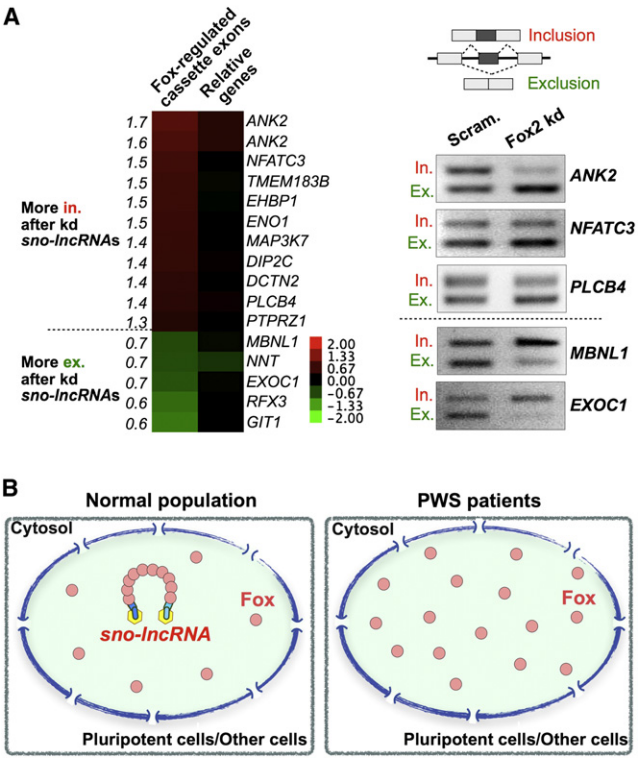

Figure 6. A Functional Connection between PWS Region *sno-lncRNA*s and Fox2

(A) Knockdown of *sno-lncRNA*s affects splicing of some Fox2-regulated cassette exons. Left, heatmap of affected Fox2-regulated cassette exons and the relative expression of their corresponding genes after knockdown of all *sno-lncRNA*s (log2 ratio). The relative change of alternative splicing of each affected cassette exon is shown on the left column of the heatmap. Note that in most cases the expression of exons is altered, while that of the genes is not. Right, cassette exons affected by knockdown of *sno-lncRNA*s (left) were randomly selected and tested for their alternative splicing after knockdown of Fox2 using semiquantitative RT-PCR. Note that knockdown of Fox2 results in more exclusion of cassette exons from *ANK2*, *NFATC3*, and *PLCB4*, while more inclusion of cassette exons from *MBNL1* and *EXOC1*.

(B) A model for the role of *sno-lncRNA*s in normal cells and in Prader-Willi Syndrome. Left, in normal pluripotent cells and other cells, high expression of the PWS region *sno-lncRNA*s leads to the sequestration of Fox family splicing regulators. One possible outcome of this sequestration could be reduced availability of Fox proteins throughout the nucleus, leading to global changes in alternative splicing regulation. Another and non-mutually exclusive outcome is that the local concentration of Fox proteins in the neighborhood of *sno-lncRNA*s is increased, leading to localized effects on alternative splicing. Right, in PWS patients, where the PWS region *sno-lncRNA*s are deleted or not expressed, Fox proteins are more uniformly distributed throughout the nucleus, resulting in altered patterns of splicing regulation during early embryonic development and adulthood.

region *sno-lncRNA*s, suggesting more similar genomewide analyses will be required to further identify more types of *sno-lncRNA*s in different cells.

The global analyses of PWS region *sno-lncRNA* knockdown revealed no dramatic changes in cellular gene expression, consistent with the observation that their lack of expression in PWS patients is not lethal. However, some Fox-regulated cassette exons showed inclusion/exclusion alterations after *sno-lncRNA* knockdown (Table S3 and Figure 6A), consistent with a model whereby the PWS region *sno-lncRNA*s act as molecular sinks. It is interesting to note that among the group of genes most affected by PWS region *sno-lncRNA* knockdown are several with a clear connection to neuronal function (Table S3). For example, Neuronal ankyrin 2 (ANK2) is expressed preferentially in the brain (Otto et al., 1991). In addition, PTPRZ1, a protein tyrosine phosphatase, binds specifically to members of the Contactin family of neural adhesion molecules which modulate oligodendrogenesis (Lamprianou et al., 2011), and GIT1 is associated with attention deficit hyperactivity disorder (Won et al., 2011). Subtle changes of splicing regulation of these cassette exons during early embryonic development and adulthood may result in an abnormal development in PWS patients.

A role as sinks for RNA-binding proteins would place PWS region *sno-lncRNA*s in a similar category with a number of other lncRNAs. For example, the *Drosophila* heat shock-induced RNA *hsr-omega* sequesters RNA binding proteins (Jolly and Lakhotia, 2006), CUG repeats sequester CUGBP1 and MBNL1 (de Haro et al., 2006; Fardaei et al., 2001), stress-induced SatIII sequesters hnRNP proteins (Valgardsdottir et al., 2008), *NEAT1* and *MALAT1* have many binding sites for TDP-43 (Polymenidou et al., 2011; Tollervey et al., 2011) and *MALAT1* binds many SR proteins and modulates alternative splicing (Tripathi et al., 2010).

Taken together, our data lead us to hypothesize that one or more of the *sno-lncRNA*s from the PWS region are functionally connected to this disease by acting as molecular sinks for Fox proteins (Figure 6B). It was recently reported that aberrant expression of Fox1 is generally associated with a variety of autisms (Voineagu et al., 2011). This may relate to PWS and to the fact that PWS critical region transcripts bind strongly to Fox2 in H9 cells (Yeo et al., 2009) and, by inference, will likewise bind strongly to Fox1 in the brain, and to both in other tissues. However, the observation that *sno-lncRNA* knockdown leads to altered expression levels of some genes (Figure 4G) leads us to emphasize that Fox protein titration may not be the only function for these *sno-lncRNA*s and more work is warranted to learn about additional protein partners and functions.

EXPERIMENTAL PROCEDURES

Cell Culture, Plasmid Construction and Cell Transfection

All human cell lines were cultured using standard protocols. Full length *sno-lncRNA2* flanked by its natural intron, splicing sites and exons was cloned into pcDNA3, and then subcloned into the EGFP reporter pZW1 (Wang et al., 2004). All deletions and mutations of snoRNA genes in *sno-lncRNA2* in the vector pcDNA3-sno-lncRNA2 were created using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). Transfection was carried out with either X-tremeGENE 9 (Roche) according to the manufacturer's protocols.

Total RNA Isolation, Human Tissue RNA Samples, RT-PCR, RT-qPCR, and Northern Blotting

Total RNAs from cultured cell lines were extracted with Trizol (Invitrogen). 20 human tissue RNA samples were purchased from Ambion. For RT-PCR, after treatment with DNase I (Ambion, DNA-free™ kit), the cDNA was transcribed with SuperScript II (Invitrogen). For qPCR, the relative expression of different sets of genes was quantified to *actin* mRNA. Northern blots were performed with Digoxigenin labeled antisense probes (DIG Northern Starter Kit, Roche).

Nuclear/Cytoplasmic RNA Fractionation, Nuclear Soluble/Insoluble RNA Fractionation and Polyadenylated/Nonpolyadenylated RNA Separation

Nuclear and cytoplasmic RNA isolation was performed as described (Chen and Carmichael, 2009). Nuclear soluble and insoluble RNA fractionation was performed with slight modifications according to Cabianca et al. (2012). Polyadenylated and nonpolyadenylated RNA separation was carried out as described (Yang et al., 2011).

RNA In Situ Hybridization and Immunofluorescence Microscopy

RNA ISH was carried out as described (Chen and Carmichael, 2009) with slight modifications. Hybridization was carried out using in vitro transcribed digoxigenin labeled antisense probes and images were taken with a Leica TCS SP5 microscope. The diameters of *sno-lncRNA*s nuclear accumulation in scramble or pooled ASOs treated cells were measured and boxplots were plotted using a custom R script. For colocalization studies, cells were fixed for 5 min in 2% formaldehyde after RNA ISH, and immunofluorescence were performed with Rabbit anti-Fox2 (1:100, Bethyl), rabbit anti-Nucleolin (1:100, Santa Cruz Biotechnology) or mouse anti-Coilin (1:100, Sigma). The nuclei were counterstained with DAPI. All experiments were repeated at least three times.

RNA/DNA FISH

Sequential RNA/DNA FISH experiments were carried out. After RNA ISH, cells were denatured and then hybridized with denatured fluorescence labeled BAC clones recognizing chr15q11.2-12 region (Empire Genomics) overnight. After hybridization and washes, image analyses were performed on single Z stacks acquired with an Olympus IX70 DeltaVision RT Deconvolution System microscope. Signals colocalization was detected in >98% double positive cells.

RNA-Protein Complex Immunoprecipitation

RNA-protein complex immunoprecipitations were carried out as described (Chen and Carmichael, 2009). Primary antibodies used in this study were rabbit anti-Fox2 (Bethyl Laboratories), rabbit anti-Fibrillarin (Abcam), and purified rabbit IgG (Sigma).

Protein-RNA Complex Immunoprecipitation by MS2-MBP

HeLa cells were transfected with indicated plasmid before harvest. Cell pellets were resuspended in IP buffer and subjected to gentle sonication and centrifuged to obtain cell extracts. The supernatant was first incubated with MS2-MBP and then amylose resin beads (NEB) at 4°C followed by thorough washes. The protein components of sno-lncRNPs were eluted from beads with IP buffer supplemented with 12 mM Maltose, and were checked by WB with different antibodies.

Antisense Oligonucleotide Treatment

Single antisense oligonucleotides (ASOs) or combined ones were introduced to PA1 cells by nucleofection (Lonza) according to the manufacturer's protocol. Renucleofection was performed to increase the knockdown efficiency at 36 hr after the first ASO treatment. Total RNAs were collected for Illumina sequencing and RT-qPCR.

RNA-Seq, Gene Expression, and Splicing Analysis

Forty and forty-seven million 1X100 single reads from two biological replicates of each sample were analyzed as described (Yang et al., 2011). Normalized gene and exon expression levels were determined in units of BPKM and correlations of gene expression among different samples were calculated with custom R scripts. A full list of gene expression is in Table S1. Altered gene expression was compared between ASO scramble and knockdown treatments with BPKM $\geq$ 1. Further criteria for scoring were that the fold change of BPKM value of knockdown versus the BPKM value of scramble must be >1.5 or <0.7 and the p-value of fold change must be <0.05 (Table S2). Validation of gene expression was carried out with RT-qPCR.

The known Fox-regulated alternative splicing events (Zhang et al., 2008) affected by knockdown of *sno-lncRNAs* were identified from the RNA-seq data using the following criteria: (1) fold change of each Fox-regulated cassette exon must be >1.3 or <0.7; (2) the p-value of fold change must be <0.05; (3) the expression of each gene was not altered. See Figure S7 for examples and Table S3 for details.

ACCESSION NUMBERS

All raw sequencing data and bigWig track files are available for download from NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE38541.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, four tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2012.07.033.

ACKNOWLEDGMENTS

We thank Y. Fang, L. Shi, and M. Bodnar for assistance in Double FISH image analysis; J. Steitz, Z. Wang, P. Boutz, and T. Chen for materials; and J. Autuoro for WT plasmid of *sno-lncRNA2*. We are grateful to B. Graveley for critical reading of the manuscript and all lab members for helpful discussion. H1 and H9 cells were obtained from the WiCell Research Institute. This work was supported by grant 2011CBA01105 from Ministry of Science and Technology of China, grants XDA01010206, 2012OHTP08, and 2012SSTP01 from Chinese Academy of Sciences, and grant 11PJ1411000 from Shanghai Municipal Science and Technology Commission to L.-L.C./L.Y.; grant 0925347 from the National Science Foundation to G.G.C. and awards from the State of Connecticut under the Connecticut Stem Cell Research Grants Program to G.G.C. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the State of Connecticut, the Department of Public Health of the State of Connecticut, or Connecticut Innovations, Inc. L.-L.C. and G.G.C. designed the project. L.-L.C. managed and supervised the project. Q.-F.Y., L.Y., Y.Z., J.-F.X., Y.-W.W., and L.-L.C. performed experiments. L.Y., G.G.C., and L.-L.C. performed the bioinformatics analysis, analyzed the data, and wrote the paper with input from all authors.

Received: April 11, 2012
Revised: June 20, 2012
Accepted: July 27, 2012
Published online: September 6, 2012

REFERENCES

Augui, S., Nora, E.P., and Heard, E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. Nat. Rev. Genet. *12*, 429–442.

Bazeley, P.S., Shepelev, V., Talebizadeh, Z., Butler, M.G., Fedorova, L., Filatov, V., and Fedorov, A. (2008). snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. Gene *408*, 172–179.

Boisvert, F.M., van Koningsbruggen, S., Navascués, J., and Lamond, A.I. (2007). The multifunctional nucleolus. Nat. Rev. Mol. Cell Biol. *8*, 574–585.

Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. (2012). A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. Cell *149*, 819–831.

Cassidy, S.B., Schwartz, S., Miller, J.L., and Driscoll, D.J. (2012). Prader-Willi syndrome. Genet. Med. *14*, 10–26.

Cavaillé, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.P., Brosius, J., and Hüttenhofer, A. (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc. Natl. Acad. Sci. USA *97*, 14311–14316.

Chen, L.L., and Carmichael, G.G. (2009). 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.

Chen, L.L., and Carmichael, G.G. (2010). Long noncoding RNAs in mammalian cells: what, where, and why? Wiley Interdiscip Rev RNA *1*, 2–21.

de Haro, M., Al-Ramahi, I., De Gouyon, B., Ukani, L., Rosa, A., Faustino, N.A., Ashizawa, T., Cooper, T.A., and Botas, J. (2006). MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of myotonic dystrophy type 1. Hum. Mol. Genet. *15*, 2138–2145.

de Smith, A.J., Purmann, C., Walters, R.G., Ellis, R.J., Holder, S.E., Van Haelst, M.M., Brady, A.F., Fairbrother, U.L., Dattani, M., Keogh, J.M., et al. (2009). A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. Hum. Mol. Genet. *18*, 3257–3265.

Duker, A.L., Ballif, B.C., Bawle, E.V., Person, R.E., Mahadevan, S., Alliman, S., Thompson, R., Traylor, R., Bejjani, B.A., Shaffer, L.G., et al. (2010). Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader-Willi syndrome. Eur. J. Hum. Genet. *18*, 1196–1201.

Fardaei, M., Larkin, K., Brook, J.D., and Hamshere, M.G. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. Nucleic Acids Res. *29*, 2766–2771.

Filipowicz, W., and Pogać ić, V. (2002). Biogenesis of small nucleolar ribonucleoproteins. Curr. Opin. Cell Biol. *14*, 319–327.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Hao, Y., Sekine, K., Kawabata, A., Nakamura, H., Ishioka, T., Ohata, H., Katayama, R., Hashimoto, C., Zhang, X., Noda, T., et al. (2004). Apollon ubiquitinates SMAC and caspase-9, and has an essential cytoprotection function. Nat. Cell Biol. *6*, 849–860.

Hirose, T., and Steitz, J.A. (2001). Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. Proc. Natl. Acad. Sci. USA *98*, 12914–12919.

Hirose, T., Shu, M.D., and Steitz, J.A. (2003). Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. Mol. Cell *12*, 113–123.

Hung, T., Wang, Y., Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings, H.M., Shah, N., Umbricht, C., Wang, P., et al. (2011). Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat. Genet. *43*, 621–629.

Jolly, C., and Lakhotia, S.C. (2006). Human sat III and Drosophila hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. Nucleic Acids Res. *34*, 5508–5514.

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., et al. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc. Natl. Acad. Sci. USA *106*, 11667–11672.

Kim, K.K., Adelstein, R.S., and Kawamoto, S. (2009). Identification of neuronal nuclei (NeuN) as Fox-3, a new member of the Fox-1 gene family of splicing factors. J. Biol. Chem. *284*, 31052–31061.

Kim, K.K., Kim, Y.C., Adelstein, R.S., and Kawamoto, S. (2011). Fox-3 and PSF interact to activate neural cell-specific alternative splicing. Nucleic Acids Res. *39*, 3064–3078.

Kiss, T. (2001). Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J. *20*, 3617–3622.

Kiss, T., and Filipowicz, W. (1995). Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. Genes Dev. *9*, 1411–1424.

Lamprianou, S., Chatzopoulou, E., Thomas, J.L., Bouyain, S., and Harroch, S. (2011). A complex between contactin-1 and the protein tyrosine phosphatase PTPRZ controls the development of oligodendrocyte precursor cells. Proc. Natl. Acad. Sci. USA *108*, 17498–17503.

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.

Nakahata, S., and Kawamoto, S. (2005). Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. Nucleic Acids Res. *33*, 2078–2089.

Ørom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., et al. (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46–58.

Otto, E., Kunimoto, M., McLaughlin, T., and Bennett, V. (1991). Isolation and characterization of cDNAs encoding human brain ankyrins reveal a family of alternatively spliced genes. J. Cell Biol. *114*, 241–253.

Polymenidou, M., Lagier-Tourenne, C., Hutt, K.R., Huelga, S.C., Moran, J., Liang, T.Y., Ling, S.C., Sun, E., Wancewicz, E., Mazur, C., et al. (2011). Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat. Neurosci. *14*, 459–468.

Rearick, D., Prakash, A., McSweeny, A., Shepard, S.S., Fedorova, L., and Fedorov, A. (2011). Critical association of ncRNA with introns. Nucleic Acids Res. *39*, 2357–2366.

Royo, H., and Cavaillé, J. (2008). Non-coding RNAs in imprinted gene clusters. Biol. Cell *100*, 149–166.

Runte, M., Hüttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001). The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. Hum. Mol. Genet. *10*, 2687–2700.

Sahoo, T., del Gaudio, D., German, J.R., Shinawi, M., Peters, S.U., Person, R.E., Garnica, A., Cheung, S.W., and Beaudet, A.L. (2008). Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat. Genet. *40*, 719–721.

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS ONE *7*, e30733.

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., 3rd, Frank, J., and Manley, J.L. (2009). Molecular architecture of the human pre-mRNA 3′ processing complex. Mol. Cell *33*, 365–376.

Skryabin, B.V., Gubar, L.V., Seeger, B., Pfeiffer, J., Handel, S., Robeck, T., Karpova, E., Rozhdestvensky, T.S., and Brosius, J. (2007). Deletion of the MBII-85 snoRNA gene cluster in mice results in postnatal growth retardation. PLoS Genet. *3*, e235.

Sokka, A.L., Mudo, G., Aaltonen, J., Belluardo, N., Lindholm, D., and Korhonen, L. (2005). Bruce/apollon promotes hippocampal neuron survival and is downregulated by kainic acid. Biochem. Biophys. Res. Commun. *338*, 729–735.

Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Zupunski, V., et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat. Neurosci. *14*, 452–458.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol. Cell *39*, 925–938.

Tycowski, K.T., Shu, M.D., and Steitz, J.A. (1993). A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. Genes Dev. *7* (7A), 1176–1190.

Underwood, J.G., Boutz, P.L., Dougherty, J.D., Stoilov, P., and Black, D.L. (2005). Homologues of the Caenorhabditis elegans Fox-1 protein are neuronal splicing regulators in mammals. Mol. Cell. Biol. *25*, 10005–10016.

Valgardsdottir, R., Chiodi, I., Giordano, M., Rossi, A., Bazzini, S., Ghigna, C., Riva, S., and Biamonti, G. (2008). Transcription of Satellite III non-coding RNAs is a general stress response in human cells. Nucleic Acids Res. *36*, 423–434.

Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J., and Geschwind, D.H. (2011). Transcriptomic

analysis of autistic brain reveals convergent molecular pathology. Nature *474*, 380–384.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol. Cell *43*, 904–914.

Wang, Z., Rolish, M.E., Yeo, G., Tung, V., Mawson, M., and Burge, C.B. (2004). Systematic identification and analysis of exonic splicing silencers. Cell *119*, 831–845.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Won, H., Mah, W., Kim, E., Kim, J.W., Hahm, E.K., Kim, M.H., Cho, S., Kim, J., Jang, H., Cho, S.C., et al. (2011). GIT1 is associated with ADHD in humans and ADHD-like behaviors in mice. Nat. Med. *17*, 566–572.

Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G., and Chen, L.L. (2011). Genomewide characterization of non-polyadenylated RNAs. Genome Biol. *12*, R16.

Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.D., and Gage, F.H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat. Struct. Mol. Biol. *16*, 130–137.

Zhang, C., Zhang, Z., Castle, J., Sun, S., Johnson, J., Krainer, A.R., and Zhang, M.Q. (2008). Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. Genes Dev. *22*, 2550–2563.