Cell

Stalled Spliceosomes Are a Signal for RNAi-Mediated Genome Defense

Phillip A. Dumesic,[1] Prashanthi Natarajan,[1] Changbin Chen,[1] Ines A. Drinnenberg,[2] Benjamin J. Schiller,[1] James Thompson,[3] James J. Moresco,[3] John R. Yates III,[3] David P. Bartel,[2] and Hiten D. Madhani[1,*]
[1]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA
[2]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA
[3]Department of Chemical Physiology, The Scripps Research Institute, La Jolla, CA 92037, USA
*Correspondence: hitenmadhani@gmail.com
http://dx.doi.org/10.1016/j.cell.2013.01.046

SUMMARY

Using the yeast *Cryptococcus neoformans*, we describe a mechanism by which transposons are initially targeted for RNAi-mediated genome defense. We show that intron-containing mRNA precursors template siRNA synthesis. We identify a *S*pliceosome-*C*oupled *A*nd *N*uclear *R*NAi (SCANR) complex required for siRNA synthesis and demonstrate that it physically associates with the spliceosome. We find that RNAi target transcripts are distinguished by suboptimal introns and abnormally high occupancy on spliceosomes. Functional investigations demonstrate that the stalling of mRNA precursors on spliceosomes is required for siRNA accumulation. Lariat debranching enzyme is also necessary for siRNA production, suggesting a requirement for processing of stalled splicing intermediates. We propose that recognition of mRNA precursors by the SCANR complex is in kinetic competition with splicing, thereby promoting siRNA production from transposon transcripts stalled on spliceosomes. Disparity in the strength of expression signals encoded by transposons versus host genes offers an avenue for the evolution of genome defense.

INTRODUCTION

RNA interference (RNAi)-related RNA silencing pathways constitute a group of small-RNA-based silencing mechanisms that antedate expansion of the eukaryotic lineage and function throughout this domain of life (Shabalina and Koonin, 2008). Enzymes required for RNA silencing are numerous and can differ between species but universally include Argonaute or PIWI clade proteins, which bind small RNAs. Some RNA silencing pathways also utilize Dicer ribonucleases, which produce small interfering RNA (siRNA) from double-stranded RNA (dsRNA) precursors, and RNA-dependent RNA polymerases, which produce dsRNA. Although RNAi-related systems perform disparate roles in different organisms—from histone modification to translational regulation—a deeply conserved and biologically critical function for these systems, observed from protists to man, is to defend genome integrity by silencing transposable elements (Behm-Ansmant et al., 2006; Cam et al., 2009; Ghildiyal and Zamore, 2009). Yet transposons occur in many families that bear little or no resemblance to each other (Malone and Hannon, 2009), raising the question of how they are recognized as nonself DNA.

One RNAi-related system that suppresses transposon mobilization is the Piwi-interacting small RNA (piRNA) pathway, best understood in *Drosophila*. piRNAs derive from specific genomic clusters of transposon-related sequences and act with Argonaute proteins of the PIWI clade to silence homologous sequences throughout the genome (Brennecke et al., 2007; Malone and Hannon, 2009). Such a mechanism constitutes an adaptive immunity to transposons, as it silences only transposon families that had previously been incorporated into a piRNA cluster (Khurana et al., 2011). These constraints raise the question of whether eukaryotes also demonstrate innate immunity to transposons, in which prior exposure to a transposon is not required for its recognition.

The processing of long dsRNA into siRNA can be viewed as an innate immune mechanism for transposon defense, capable of recognizing even novel transposons by virtue of their tendency to generate dsRNA. For instance, transposons can produce dsRNA by mobilizing into an existing transcriptional unit or by virtue of transposon-encoded inverted repeats and internal antisense promoters; such dsRNAs template the production of repressive endogenous siRNA in a manner that requires Dicer (Conley et al., 2008; Drinnenberg et al., 2009; Ghildiyal et al., 2008; Sijen and Plasterk, 2003; Yang and Kazazian, 2006). Mutations that block exogenous RNAi, which is triggered by long dsRNA, concomitantly increase endogenous transposon mobilization, providing additional evidence for the role of dsRNA processing in transposon recognition (Ketting et al., 1999; Tabara et al., 1999). Another class of RNAi-related system potentially involved in innate transposon immunity is thought to have evolved to recognize unusual DNA arrangements. The quelling pathway of *Neurospora crassa* targets repetitive transgene arrays (Lee et al., 2010; Nolan et al., 2008), whereas meiotic silencing of unpaired DNA (MSUD) mechanisms silence transgenes that lack a partner during homolog pairing in meiosis I (Kelly and Aramayo, 2007). Unlike

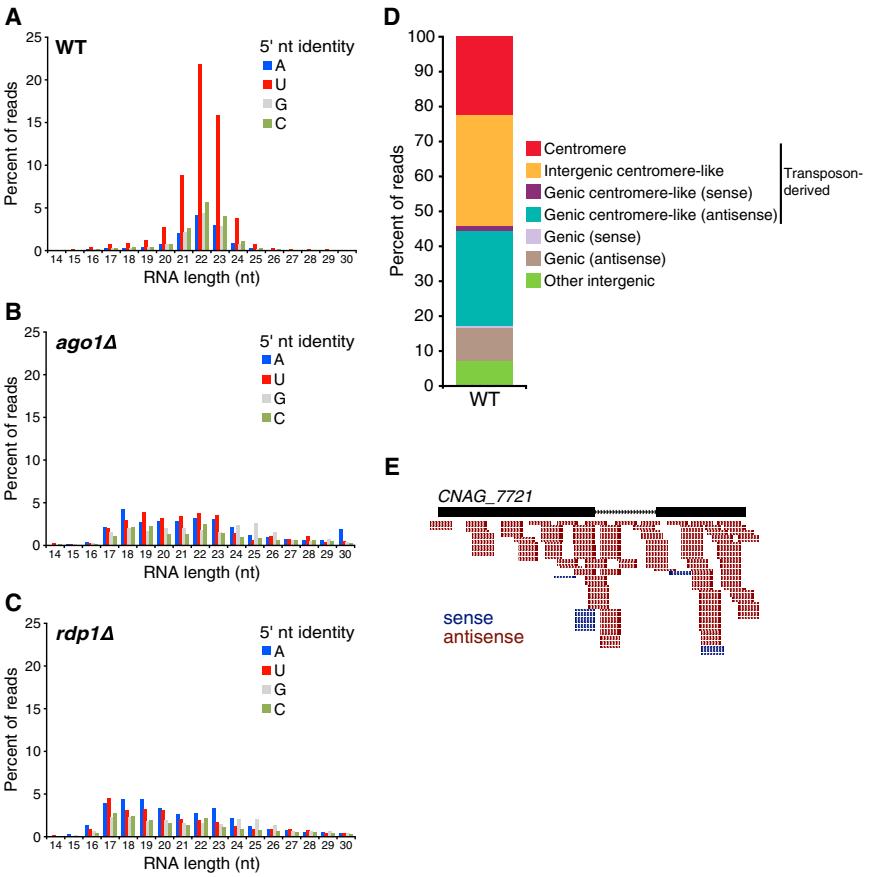


Figure 1. Endogenous siRNA of *C. neoformans*

(A–C) Read counts for siRNAs based on length and 5′ nucleotide identity.

(D) Genomic mapping of siRNA sequences.

(E) Density plot of siRNAs mapping to an RNAi target locus, *CNAG_7721*, which comprises two exons and one intron.

See also Tables S1 and S2.

mechanisms of innate transposon immunity that rely on dsRNA recognition, quelling and MSUD can silence loci that do not naturally produce dsRNA, and both require RNA-dependent RNA polymerases (Cogoni and Macino, 1999; She et al., 2009; Shiu et al., 2001). Because transposons are particularly likely to form tandem arrays or occur asymmetrically on homologs during meiosis, quelling and MSUD may suppress transposon mobilization, but their underlying mechanisms remain poorly understood.

The human pathogenic yeast *Cryptococcus neoformans* offers a genetically tractable model system for approaching the mechanisms of small-RNA-mediated transposon suppression. This organism displays an active RNAi pathway comprising one Argonaute clade member (Ago1), two redundant Dicer orthologs (Dcr1/2), and an RNA-dependent RNA polymerase ortholog (Rdp1). These factors play a key role in defending the *C. neoformans* genome: null mutations in their corresponding genes result in increased transposon expression, transposon mobilization, and transposon-induced drug resistance mutations (Janbon et al., 2010; Wang et al., 2010). In this paper, we investigate the mechanism of siRNA biogenesis in *C. neoformans*. Six observations indicate a key role for introns and the spliceosome in this process. First, small RNA analysis reveals that unspliced messenger RNA (mRNA) precursors are preferred substrates for siRNA production. Second, we describe a nuclear RNA-dependent RNA polymerase complex required for siRNA production, termed SCANR (*S*pliceosome-*C*oupled *A*nd *N*uclear *R*NAi), and find that it physically associates with the spliceosome. Third, we observe that RNAi target transcripts encode suboptimal splicing signals and exhibit unusually high accumulation on spliceosomes. Fourth, we find that deletion of introns from a strong RNAi target blocks the accumulation of siRNA corresponding to this transcript. Fifth, we find that experimental stalling of a pre-mRNA's splicing dramatically increases its siRNA production in a manner that requires entry into the splicing pathway. Sixth, we demonstrate that the lariat debranching enzyme is required for siRNA synthesis. These results indicate that stalled spliceosomes are a signal for RNAi and that splicing intermediates may be a favored substrate. We propose that a competition between the recognition of spliceosome-associated mRNA precursors by SCANR and the completion of their ongoing splicing plays an important role in specifying sequences from which repressive siRNA is produced.

RESULTS

Endogenous siRNA of *C. neoformans* Targets Transposon mRNA Precursors

To investigate the mechanism of small RNA biogenesis in *C. neoformans* genome defense, we used high-throughput sequencing to identify the siRNAs in haploid, vegetatively growing cells. Wild-type cells produced 21–23 nt siRNAs with a strong preference for a uridine residue at their 5′ termini, a characteristic feature of Argonaute-bound siRNAs in other fungal systems (Figure 1A) (Drinnenberg et al., 2009). These siRNAs were absent from cells lacking Ago1 or Rdp1 (Figures 1B and 1C). siRNAs mapped most prominently to transposons: 22% of reads mapped to centromeres, which are composed primarily of transposons and transposon remnants (Loftus et al., 2005), and 32% of reads mapped to intergenic regions that have sequence similarity to centromeres (Figure 1D). An additional 39% of siRNA reads mapped to mRNA-encoding genes, many of which had sequence similarity to centromeres. Genic siRNAs displayed a strand bias, as 95% of reads had a polarity

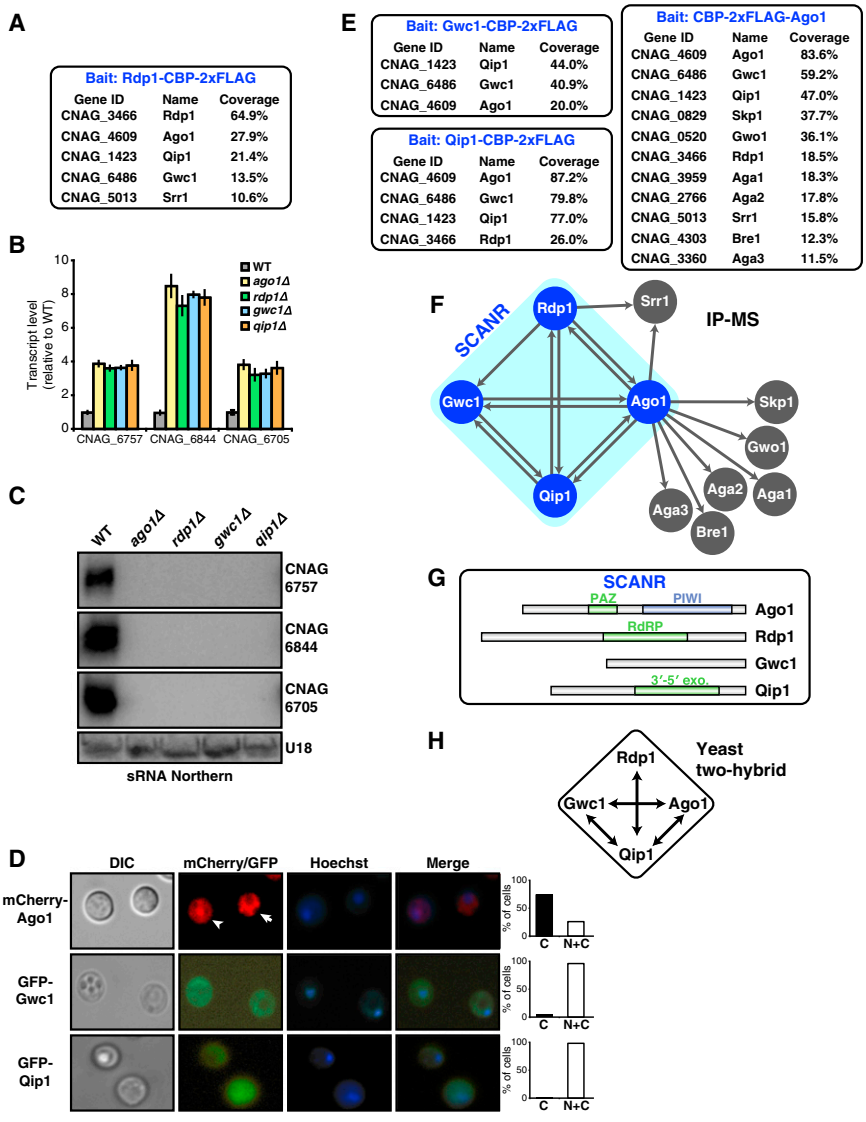


Figure 2. Functional Characterization of Rdp1-Associated Proteins

(A) Proteins associated with Rdp1-CBP-2xFLAG by tandem affinity purification. Likely contaminants and proteins with <10% sequence coverage have been excluded.

(B) Transcript levels of three RNAi target genes, assessed by RT-qPCR and normalized to 18S rRNA levels. Error bars represent SEM.

(C) sRNA Northern blot. Loading control: U18 snoRNA.

(D) Localization of Rdp1-associated proteins. Proteins were fused to fluorescent protein tags and detected by epifluorescence microscopy. Nuclei were labeled using Hoechst dye. Percentage of cells exhibiting cytoplasmic localization versus dual nuclear and cytoplasmic localization is indicated. White arrow: cytoplasmic localization; white arrowhead: nuclear and cytoplasmic localization.

(E) Proteins that interact with the Rdp1-associated proteins Gwc1, Qip1, and Ago1.

(F) Protein interaction network of Rdp1-associated proteins. The purification results of (A) and (E) are represented graphically.

(G) Predicted protein domains of SCANR subunits.

(H) Protein-protein interactions among SCANR components, as assessed by yeast two-hybrid assay.

See also Figure S1 and Table S3.

opposite that of the predicted transcript at their corresponding loci. These findings indicate that the previously described broad targeting of transposon transcripts by siRNA that occurs during mating is also a feature of vegetative growth in *C. neoformans* (Wang et al., 2010).

Transcripts targeted by RNAi corresponded to loci on each of the 14 *C. neoformans* chromosomes (e.g., Figure 1E). These RNAi targets were not enriched for convergently transcribed genes, suggesting that dsRNA generated by bidirectional transcription is not a major driver of siRNA generation (Table S1 available online). Strikingly, siRNA reads in genes mapped not only to exons but also to introns; more than 15% of these siRNAs contained intronic sequence, whereas introns constitute 18% of all genic sequences. Moreover, the 50 genes targeted by the greatest number of siRNA reads exhibited not only intronic siRNAs (in 45 cases) but also siRNAs that spanned intron-exon junctions (in 42 cases) (Table S2). siRNA reads spanning exon-exon junctions were also observed within this gene set at a frequency only 2.45-fold that of intron-exon junction reads, despite the fact that the vast majority of cellular RNAs are expected to be fully spliced (Table S2). The sequence features of siRNAs, together with their dependence on Rdp1, are consistent with a model in which Rdp1 acts preferentially upon intron-containing mRNA precursors to generate dsRNA, which is processed into siRNA to effect silencing of transposons and other RNAi targets. This model is further developed below.

Identification and Characterization of Proteins Associated with RNA-Dependent RNA Polymerase

To gain further insights into how the RNAi machinery might be targeted to mRNA precursors in the nucleus, we used tandem affinity purification and mass spectrometry to identify proteins associated with Rdp1, the only component of the *C. neoformans* RNAi machinery that has been reported to be nuclear (Wang et al., 2010). Purification of Rdp1-CBP-2xFLAG, expressed from its endogenous locus, yielded Rdp1 itself as well as four additional proteins: Ago1 and three proteins we named Qip1, Gwc1, and Srr1 (Figure 2A). Qip1 is an ortholog of QIP, an *N. crassa* exonuclease that binds Argonaute and degrades the passenger strand of siRNA duplexes (Maiti et al., 2007). Gwc1 (GW-containing) contains five GW/WG dipeptides—a motif commonly found in Argonaute-binding

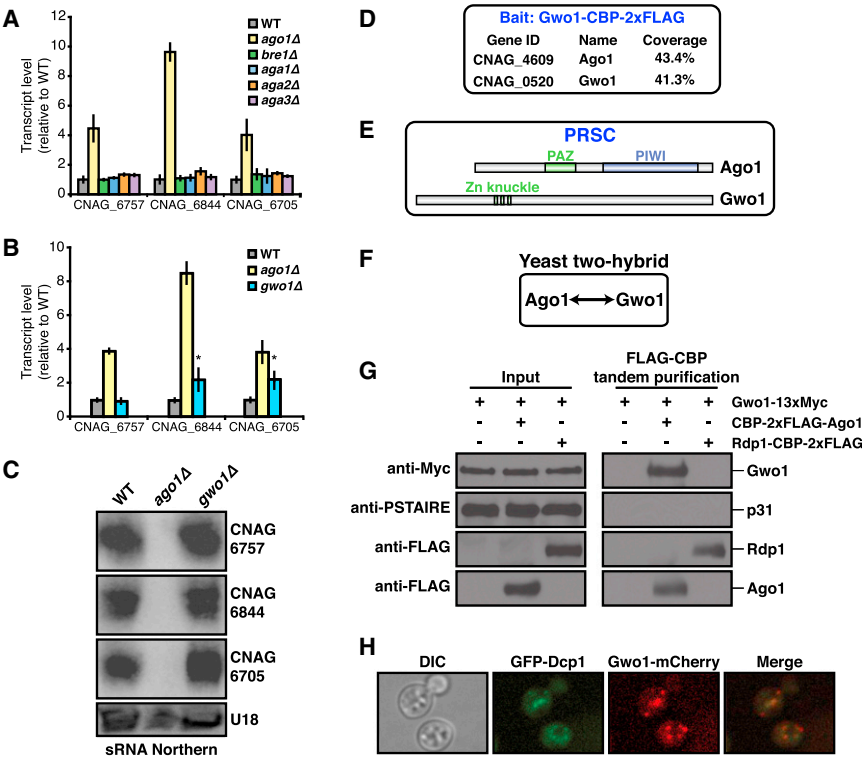


Figure 3. PRSC and SCANR Subunits Are Physically and Functionally Distinct

(A and B) Transcript levels of three RNAi target genes, assessed by RT-qPCR and normalized to 18S rRNA levels. Error bars represent SEM. *$p < 0.05$ by Student's t test.

(C) sRNA Northern blot. Loading control: U18 snoRNA.

(D) Proteins associated with Gwo1-CBP-2xFLAG. Assayed as in Figure 2A.

(E) Predicted protein domains of PRSC subunits.

(F) Protein-protein interactions among PRSC components, as assessed by yeast two-hybrid assay. No interactions were detected between Gwo1 and any exclusive member of SCANR.

(G) Coimmunoprecipitation of Gwo1 with Ago1, but not with Rdp1. Strains expressing Gwo1-13xMyc and CBP-2xFLAG-Ago1 or Rdp1-CBP-2xFLAG were subjected to FLAG-CBP tandem purification. Input and purified material were analyzed by immunoblot using anti-FLAG, anti-Myc, or anti-PSTAIRE, which detects the negative control protein p31.

(H) Colocalization of PRSC component Gwo1 with the P body marker Dcp1. Unfixed cells were examined after incubation in yeast nitrogen base (YNB) media.

See also Figure S1 and Table S3.

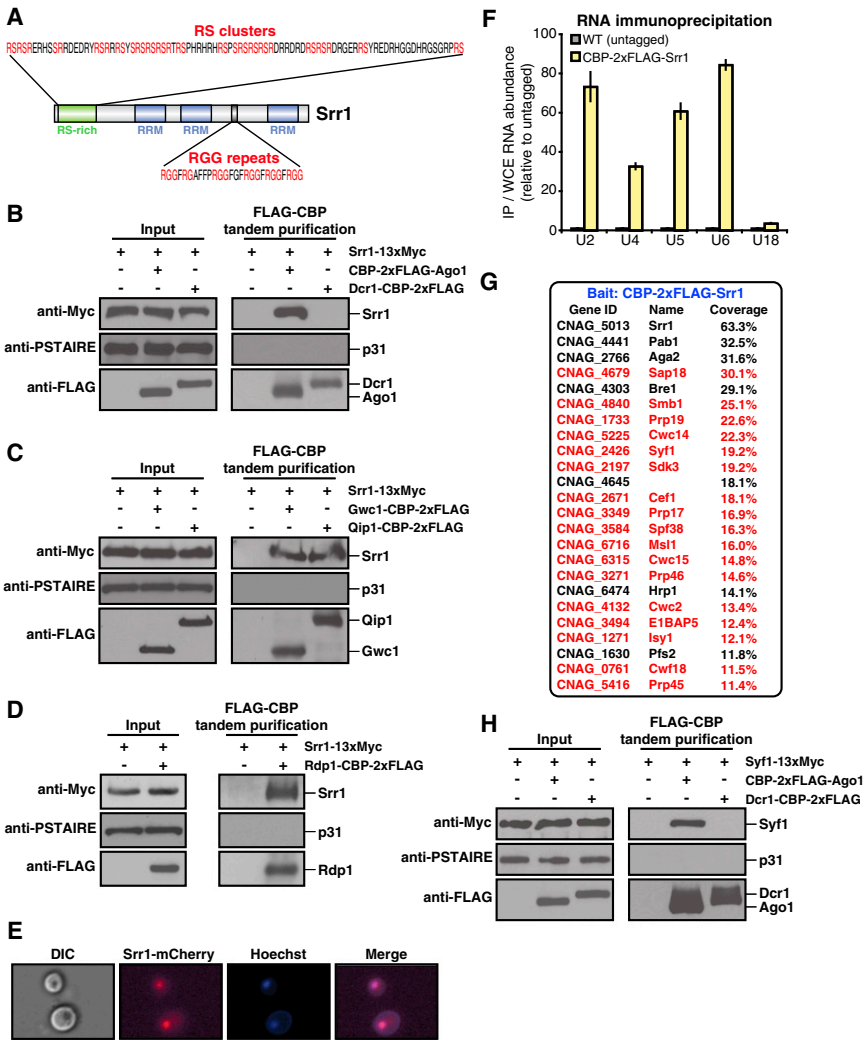
proteins (El-Shami et al., 2007)—but no other obvious domain or homolog. Srr1 (Serine/Arginine-rich) has no clear ortholog in other organisms, but, intriguingly, its domain structure resembles those of several mammalian splicing factors. In particular, Srr1 contains 30 RS/SR dipeptides and three RNA recognition motifs.

To assess the potential function of each protein that copurified with Rdp1, we attempted to create deletion mutations in their corresponding genes. We were successful for Ago1, Qip1, and Gwc1 but were unsuccessful for Srr1, suggesting that it may be essential for viability. Strains lacking Ago1, Rdp1, Gwc1, or Qip1 exhibited increased levels of three transcripts highly targeted by siRNA in our sequencing experiments: *CNAG_6757*, an unannotated gene with homology to a transposon; *CNAG_6844*, a RecQ helicase; and *CNAG_6705*, an unannotated open reading frame (ORF) (Figure 2B). The same strains also displayed a loss of siRNA corresponding to these transcripts, indicating a functional requirement for Ago1, Rdp1, Gwc1, and Qip1 in RNAi (Figure 2C).

The isolation of Ago1 as an Rdp1-associated protein was surprising because it had been reported to localize exclusively to cytoplasmic P bodies (Wang et al., 2010). However, we found that, although an mCherry-Ago1 fusion protein was cytoplasmic in most cells, a substantial fraction of cells (26%) displayed localization in both the nucleus and cytoplasm (Figure 2D). The cytoplasmic signal of mCherry-Ago1 localized both diffusely and in foci corresponding to P bodies (Figure S1A). GFP-Gwc1 and GFP-Qip1 fusion proteins displayed a nuclear and cytoplasmic pattern in all cells, which is consistent with a fraction of these proteins associating with Rdp1 in the nucleus (Figure 2D).

To characterize the physical relationships among Rdp1-associated proteins, we performed tandem affinity purifications of Gwc1, Qip1, and Ago1. Protein identification by mass spectrometry revealed a dense network of interactions among these proteins, which is consistent with the existence of a protein complex (Figures 2E and 2F). Gwc1 and Qip1 purifications yielded only proteins that were also identified in an Rdp1 purification, whereas an Ago1 purification yielded all of these proteins as well as six additional proteins, which we named Skp1, Gwo1, Aga1, Aga2, Bre1, and Aga3. These results can be most simply explained by proposing that Rdp1, Gwc1, Ago1, and Qip1 form a protein complex required for siRNA production, whereas Ago1 participates additionally in at least one other protein complex. For reasons described below, we refer to the Rdp1 complex as SCANR (Figure 2G). We note that the dual nuclear and cytoplasmic localization of Ago1, Gwc1, and Qip1 suggests the existence of a cytoplasmic subcomplex that lacks Rdp1. Yeast two-hybrid experiments using each SCANR component as bait and prey revealed interactions involving every subunit, further supporting the associations identified by tandem affinity purification (Figure 2H and Table S3).

We tested whether any of the proteins that copurified with Ago1, but not with any other SCANR protein (Skp1, Gwo1, Aga1, Aga2, Bre1, and Aga3), are required for silencing of mRNAs targeted by RNAi. We were able to delete each corresponding gene except *SKP1*, which may be essential for viability, as it is in *S. cerevisiae* (Connelly and Hieter, 1996). Of the five mutants, only the *gwo1Δ* strain exhibited derepression of RNAi target transcript levels, albeit to a lesser extent than that of



Gene ID	Name	Coverage
CNAG_5013	Srr1	63.3%
CNAG_4441	Pab1	32.5%
CNAG_2766	Aga2	31.6%
CNAG_4679	Sap18	30.1%
CNAG_4303	Bre1	29.1%
CNAG_4840	Smb1	25.1%
CNAG_1733	Prp19	22.6%
CNAG_5225	Cwc14	22.3%
CNAG_2426	Syf1	19.2%
CNAG_2197	Sdk3	19.2%
CNAG_4645		18.1%
CNAG_2671	Cef1	18.1%
CNAG_3349	Prp17	16.9%
CNAG_3584	Spf38	16.3%
CNAG_6716	Msl1	16.0%
CNAG_6315	Cwc15	14.8%
CNAG_3271	Prp46	14.6%
CNAG_6474	Hrp1	14.1%
CNAG_4132	Cwc2	13.4%
CNAG_3494	E1BAP5	12.4%
CNAG_1271	Isy1	12.1%
CNAG_1630	Pfs2	11.8%
CNAG_0761	Cwf18	11.5%
CNAG_5416	Prp45	11.4%

Figure 4. SCANR Physically Associates with Srr1 and the Spliceosome

(A) Predicted domains of Srr1.

(B–D) Coimmunoprecipitation of Srr1 with SCANR components Ago1 (B), Gwc1 (C), Qip1 (C), and Rdp1 (D), but not with Dcr1 (B). Assayed as in Figure 3G.

(E) Localization of Srr1. An Srr1-mCherry fusion protein was detected in fixed cells. Nuclei were stained using Hoechst dye.

(F) Interaction of Srr1 and spliceosomal snRNAs. Levels of individual RNAs coimmunoprecipitated with Srr1 were assessed by RT-qPCR and normalized to their abundance in whole-cell extract; transcript level is relative to that obtained in purifications from wild-type (untagged) lysates. Error bars represent SD.

(G) Proteins associated with CBP-2xFLAG-Srr1. Genes were named based on *S. cerevisiae* orthologs or, in the absence of one, based on metazoan orthologs. Known spliceosome components are colored in red.

(H) Coimmunoprecipitation of the spliceosome component Syf1 with Ago1, but not with Dcr1. Assayed as in Figure 3G.

See also Figure S2.

strains lacking SCANR components (Figures 3A and 3B). Loss of Gwo1 did not affect siRNA accumulation, however, indicating that Gwo1 is not required for siRNA biogenesis (Figure 3C). Tandem affinity purification of Gwo1, followed by mass spectrometry, identified only Gwo1 itself and Ago1, suggesting that these two factors form a protein complex distinct from SCANR (Figures 3D and 3E). Indeed, yeast two-hybrid analysis detected an interaction between Gwo1 and Ago1, but not between Gwo1 and the SCANR components Rdp1, Qip1, or Gwc1 (Figure 3F and Table S3). Furthermore, analytical tandem affinity purification of Ago1 yielded Gwo1, whereas purification of the SCANR component Rdp1 did not yield Gwo1, as assessed by immunoblotting (Figure 3G). Finally, localization experiments indicated that a GFP-Gwo1 fusion protein resided in cytoplasmic foci that colocalized with Ago1 foci and with the P body marker Dcp1 (Figures 3H and S1B). We therefore refer to the Ago1-Gwo1 complex as PRSC (*P*-body-associated *R*NA *S*ilencing *C*omplex) and conclude that it is distinct from the nuclear SCANR complex. Investigation of the function of PRSC is ongoing.

SCANR Associates with the Spliceosome

Given that unspliced mRNA precursors appear to be preferred substrates for dsRNA synthesis, we were intrigued by the fact that Srr1, which was identified by mass spectrometry in purifications of two SCANR components, resembles mammalian splicing regulators (Figure 4A). To confirm a physical interaction between SCANR and Srr1, we performed analytical tandem affinity purifications of SCANR components and used immunoblotting to detect associated proteins. Purifications of each SCANR subunit yielded Srr1 (Figures 4B–4D). In contrast, purification of the RNAi factor Dcr1, which has been reported to localize to P bodies (Wang et al., 2010), did not yield Srr1, nor was an unrelated control protein, p31, detected in any purification. Consistent with its SCANR association, an Srr1-mCherry fusion protein displayed a nuclear localization (Figure 4E).

Given its domain structure, we tested whether Srr1 physically associates with spliceosomal snRNAs by immunoprecipitating Srr1 and detecting bound RNAs by RT-qPCR. All four annotated *C. neoformans* spliceosomal snRNAs (U2, U4, U5, and U6), but not a control small nucleolar RNA (snoRNA) (U18), associated with Srr1 (Figure 4F). We next identified proteins associated with Srr1 using tandem affinity purification and mass spectrometry. Remarkably, of the 23 proteins that copurified with Srr1, 17 were orthologs of known spliceosomal proteins (indicated in red, Figure 4G). These included components of spliceosomal snRNPs, such as Smb1, Prp4, and Msl1, and other spliceosome-associated complexes, such as Prp19, Syf1, Cef1,

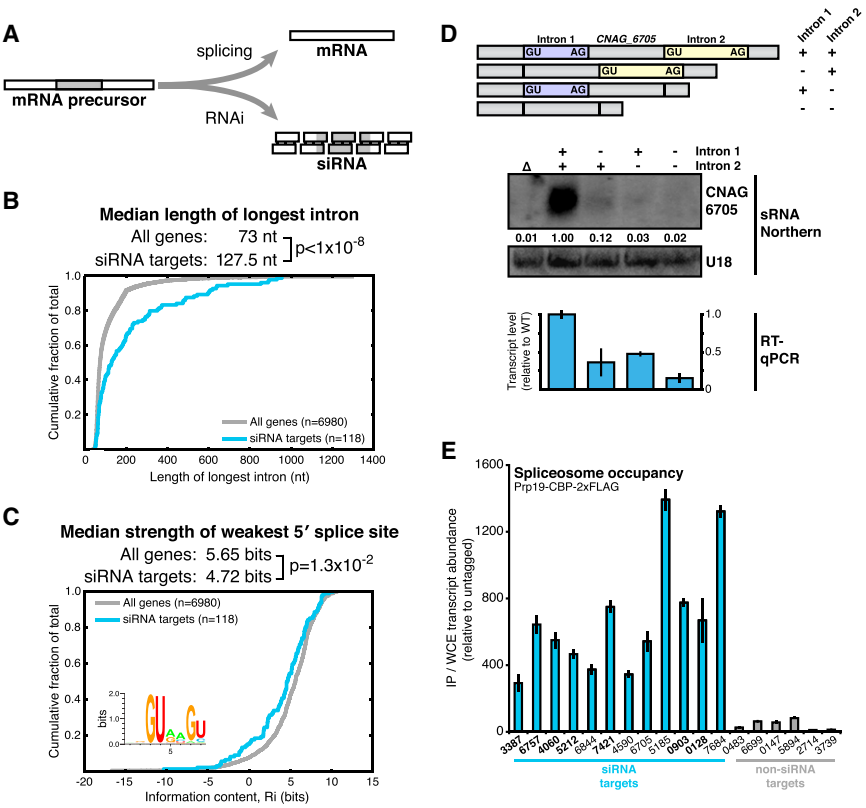


Figure 5. Transcripts Targeted by RNAi Display Suboptimal Splicing Features and Are Stalled on Spliceosomes In Vivo

(A) Model of competing pathways for mRNA precursor maturation.

(B) Analysis of longest intron lengths of siRNA targets (>150 siRNA reads) versus all genes. Cumulative distribution functions of longest intron length for these two gene classes are plotted.

(C) Analysis of weakest 5′ splice sites of siRNA targets (>150 siRNA reads) versus all genes. Cumulative distribution functions of weakest 5′ splice site strengths for these two gene classes are plotted. Inset: sequence logo generated from all *C. neoformans* 5′ splice sites.

(D) Effect of intron deletion on the production of siRNA from the *CNAG_6705* transcript. Above: RNA was isolated from cells that expressed, from a *GAL7* promoter, either wild-type *CNAG_6705* or a mutated form of the gene in which one or both introns were deleted. Below: the transcript level of each *CNAG_6705* mutant was measured by RT-qPCR and normalized to levels of actin RNA. Error bars represent SD.

(E) Association of siRNA target transcripts and non-siRNA target transcripts with the spliceosome. Spliceosomes tagged with Prp19-CBP-2xFLAG were immunoprecipitated, and copurified RNAs were detected by RT-qPCR. Levels of individual RNAs were normalized to their abundance in whole-cell extract; IP/WCE values are relative to those of purifications from wild-type (untagged) lysates. Error bars represent SD. Bold: transposon-related transcripts.

See also Figure S3.

Cwc2, and Isy1 of the NineTeen complex (NTC). Detection of SCANR components among Srr1-associated proteins was below our 10% coverage threshold, which is consistent with the lower sensitivity of this assay relative to immunoblotting. These data establish Srr1 as a nuclear protein that interacts with spliceosomal complexes.

The observed interaction between Srr1 and RNAi factors of SCANR, together with the associations between Srr1 and the spliceosome, suggested that SCANR might physically associate with the spliceosome but too weakly to be detected by mass spectrometry methods. To test this hypothesis, we determined whether Ago1 and Rdp1 physically interact with the NTC component Syf1, which is part of the spliceosome at multiple stages of its assembly and during the catalytic steps of splicing (Hogg et al., 2010). Using immunoblotting, we detected Syf1 in tandem affinity purifications of Ago1 or Rdp1, but not upon purification of Dcr1, which is not a SCANR component (Figures 4H and S2). These findings indicate that SCANR physically associates with at least a subset of spliceosomal complexes marked by Syf1.

RNAi Target Transcripts Display High Spliceosome Occupancy

Our finding that mRNA precursors are a substrate not only for splicing but also for siRNA production (Figure 1) led us to hypothesize that the spliceosome and RNAi may compete for their substrates in some way (Figure 5A). In this scenario, the slow splicing of a transcript would, by virtue of increasing its time spent as an mRNA precursor, increase its availability as a substrate for the RNAi pathway, thereby biasing RNAi targeting toward poorly spliced genes.

An important prediction of such a model is that genes targeted by RNAi should encode inefficiently spliced transcripts. As an initial test of this idea, we examined intron features likely to affect splicing efficiency. Unlike *S. cerevisiae*, in which most genes have no introns and genes with introns typically have only one, *C. neoformans* genes average five introns per gene, and it is thought that nearly all, if not all, genes harbor introns (Loftus et al., 2005). Furthermore, whereas *S. cerevisiae* intronic splicing signals conform closely to consensus, *C. neoformans* splice sites are considerably more degenerate, and in this respect they resemble those of metazoans (Irimia et al., 2007; Loftus et al., 2005). Although *C. neoformans* intron splice sites possess generally low information content, intron size is highly constrained: intron lengths are tightly distributed around a mode of 52 nt, and evolutionary studies of the *Cryptococcus* species complex suggest that an optimal intron size is under selection (Hughes et al., 2008).

We examined the aforementioned splicing features of RNAi target and nontarget transcripts to determine whether the RNAi pathway acts preferentially on transcripts with suboptimal introns. The vast majority of both gene classes contained introns—RNAi target transcripts encode 3.9 introns per gene

on average, as compared to 5.1 for all *C. neoformans* genes. Because a single poorly spliced intron in a multi-intron transcript would, in principle, be sufficient to stall maturation of the entire transcript, we classified each gene based on the properties of its weakest intron with respect to a given feature. We observed that the longest annotated introns of RNAi target genes were, on average, considerably longer than those of all genes (Figure 5B; $p < 1 \times 10^{-8}$ by two-tailed Kolmogorov-Smirnov [K-S] test). In contrast, the strengths of the weakest 5′ splice sites of RNAi target genes were, on average, only modestly weaker than those of all *C. neoformans* genes (Figure 5C; $p = 0.013$ by two-tailed K-S test). Because the 5′ splice site is critical for initiation of spliceosome assembly, whereas experimentally increasing intron length can, depending on the site, block later steps of the splicing pathway (Cellini et al., 1986; Chua and Reed, 2001), these findings raise the possibility that siRNA target transcripts enter the splicing pathway but become stalled at a later stage. Stalled splicing may thereby provide an opportunity for the spliceosome-associated SCANR complex to target mRNA precursors for RNAi.

As an initial test of this hypothesis, we examined whether causing an RNAi target transcript to bypass splicing would reduce its ability to template the production of siRNA. To do so, we generated mutants of an RNAi target transcript, *CNAG_6705*, in which its two introns were deleted, singly or together. Each of these mutations dramatically reduced the accumulation of siRNA corresponding to *CNAG_6705* (Figure 5D). Intron deletion also decreased steady-state levels of the *CNAG_6705* transcript, but this effect appeared insufficient to explain the loss of *CNAG_6705* siRNA caused by the mutations (Figure 5D). These results are consistent with a model—tested further below—in which spliceosome engagement is important for siRNA production.

To examine more directly the interactions between endogenous RNAi target transcripts and the splicing machinery, we sought to measure the in vivo occupancy of mRNA precursors on the spliceosome. We immunoprecipitated NTC-containing spliceosomal complexes using an epitope-tagged Prp19 and examined levels of associated transcripts by RT-qPCR, normalizing the data for total transcript levels. Strikingly, the 12 transcripts most highly targeted by siRNA exhibited dramatically greater spliceosome occupancy than did six non-RNAi target genes that spanned a broad range of expression level, intron number, and gene size (Figure 5E). These 12 RNAi targets include seven transposon-related genes (indicated in bold, Figure 5E), whereas the non-RNAi targets include genes encoding actin (*CNAG_0483*) and GAPDH (*CNAG_6699*). We found that the accumulation of RNAi target transcripts on spliceosomes was maintained in the context of an *rdp1Δ* strain, demonstrating that this property is not caused by siRNAs themselves (Figure S3). Thus, transcripts targeted by RNAi exhibit intronic sequence features predictive of poor splicing and accumulate abnormally on spliceosomes in vivo.

Mutations that Stall Splicing Trigger RNAi

Our finding that RNAi target transcripts accumulate on spliceosomes led us to hypothesize that these transcripts proceed inefficiently through the splicing pathway, during which some fraction is redirected to the RNAi pathway and converted into siRNA. This hypothesis predicts that the entry of a transcript into the RNAi pathway would be improved by intronic mutations that stall its splicing at an intermediate stage. To test this prediction, we introduced intronic mutations into a single-copy, endogenous gene that is very weakly targeted by RNAi: *CNAG_7888* (Figure S4A). *CNAG_7888* contains two predicted introns whose locations we verified by complementary DNA sequencing. To facilitate detection of siRNA generated from the *CNAG_7888* locus by Northern hybridization, its promoter was replaced with the strong *GAL7* promoter. We assessed two classes of intronic mutations: 5′ splice site mutations, which we predicted to block splicing prior to spliceosome assembly, and 3′ splice site mutations, which we predicted to allow spliceosome assembly but block the second catalytic step of splicing.

Cells expressing wild-type *CNAG_7888* exhibited a very low level of siRNA, comparable to the background observed in cells in which the *CNAG_7888* locus was deleted (Figure 6A). No change in siRNA levels was detected in the context of *CNAG_7888* alleles containing 5′ splice site mutations of either intron. In contrast, a 3′ splice site mutation—specifically, that of intron 2—resulted in dramatically increased siRNA production from *CNAG_7888*. This siRNA production required Rdp1, confirming that it represented entry of *CNAG_7888* into the RNAi pathway (Figure S4B). The inability of other *CNAG_7888* mutant alleles to promote siRNA generation could not be explained by defects in their expression, as all mutants showed similar transcript levels (Figure S4C). Additionally, it was not the case that the sequence of intron 2 is uniquely capable of promoting RNAi, as a *CNAG_7888* construct in which the sequences of introns 1 and 2 had been exchanged similarly triggered siRNA synthesis only when the 3′ splice site of the downstream intron was mutated (Figure S4D). Evidently, the position of the mutated intron in *CNAG_7888* (or perhaps its contiguous exonic sequence) influences its ability to stimulate siRNA production. The nature of the 3′ splice site mutation is unimportant: multiple, distinct 3′ splice site mutations in the second intron stalled *CNAG_7888* splicing at the second catalytic step, as assessed by primer extension detection of lariat intermediate, and each strongly triggered siRNA production (Figure 6B).

Based on our finding that RNAi target transcripts are enriched for introns predicted to be longer than optimal (Figure 5B), we hypothesized that introns containing an increased distance from the branchpoint adenosine to the 3′ splice site would promote RNAi because such introns stall splicing in other systems (Cellini et al., 1986; Chua and Reed, 2001). We generated alleles of *CNAG_7888* that contained 75 or 100 nt insertions of adenosine-free sequence between the mapped branchpoint and 3′ splice site of intron 2 and then examined their splicing efficiency and siRNA production. As predicted, both of these insertion mutations reduced intron 2 splicing efficiency, although not as severely as did mutation of the 3′ splice site, as assessed by level of lariat intermediate (Figure S4E). Accordingly, both insertion mutations triggered siRNA production from *CNAG_7888* but to a lesser extent than that triggered by a 3′ splice site mutation. This inverse correspondence between splicing efficiency and siRNA accumulation across a range of splicing efficiencies is consistent with a kinetic competition between splicing and

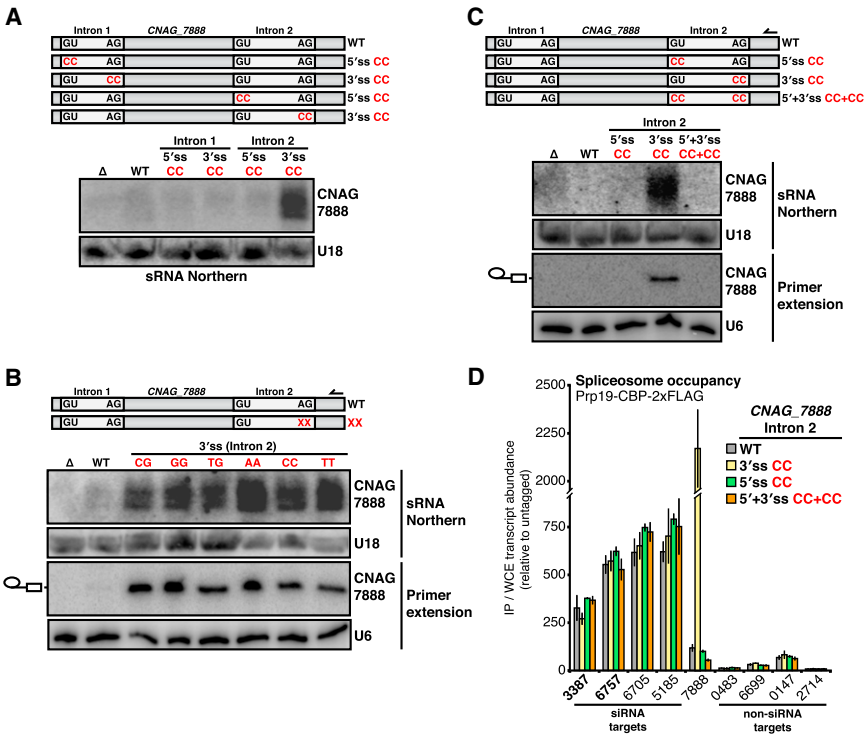


Figure 6. Stalled Spliceosomes Promote RNAi in a Manner that Depends on Entry into the Splicing Pathway

(A) Effect of perturbed splicing on the production of siRNA from the *CNAG_7888* transcript. RNA was isolated from cells that expressed, from a *GAL7* promoter, either wild-type *CNAG_7888* or a mutated form of the gene in which a single splice site was mutated.

(B) Effect of 3′ splice site mutations on siRNA production from and splicing of the *CNAG_7888* transcript. RNA was isolated from cells that expressed, from a *GAL7* promoter, either wild-type *CNAG_7888* or a mutated form of the gene in which the 3′ splice site of intron 2 was mutated. siRNA production was assessed by riboprobe hybridization, and splicing was assessed by primer extension using a labeled primer complementary to *CNAG_7888* exon 3. Loading control: U6 snRNA primer extension product.

(C) Spliceosome dependence of siRNA accumulation triggered by 3′ splice site mutations. RNA was isolated from cells that expressed, from a *GAL7* promoter, either wild-type *CNAG_7888* or a mutated form of the gene in which the 5′ and 3′ splice sites of intron 2 were mutated, singly or together, and assayed as in (B).

(D) Association of the spliceosome with *CNAG_7888* transcripts containing intron 2 splice site mutations, assayed as in Figure 5E. Error bars represent SD.

See also Figure S4.

siRNA production and indicates that multiple types of suboptimal introns can promote RNAi.

Entry into the Splicing Pathway Is Required for RNAi

Our finding that the suboptimal introns that promote RNAi tend to cause stalling of the spliceosome, together with our observation that RNAi target transcripts accumulate on spliceosomes, led us to hypothesize that spliceosome assembly at an intron is required for that intron to promote RNAi. This hypothesis predicts that a 5′ splice site mutation, which disrupts the initial recognition of an intron by the spliceosome, should suppress the RNAi-promoting effects of 3′ splice site mutations in *CNAG_7888*. Strikingly, a 5′ splice site mutation of intron 2, which caused no effect on RNAi by itself, fully suppressed the siRNA production triggered by a 3′ splice site mutation (Figure 6C). As expected, the 5′ splice site mutation also suppressed the accumulation of intron 2 lariat intermediate observed in a 3′ splice site mutant.

To confirm that 3′ splice site mutations stall spliceosomes and that this effect is eliminated by a 5′ splice site mutation, we employed the spliceosome occupancy assay described above. As expected based on its exceedingly weak production of siRNA, wild-type *CNAG_7888* exhibited low spliceosome occupancy, similar to that of non-RNAi target transcripts (Figure 6D). A 5′ splice site mutation of intron 2 caused no change in spliceosome occupancy, whereas a 3′ splice site mutation dramatically increased spliceosome occupancy, which is consistent with our finding that this mutation blocks the second catalytic step of splicing. In accord with the well-established role of the 5′ splice site sequence in mediating entry into the splicing pathway, a 5′ + 3′ double splice site mutant displayed low spliceosome occupancy. The correlation between spliceosome occupancy and siRNA levels in this series of *CNAG_7888* mutant transcripts mirrors our observation that RNAi target transcripts generally display greater spliceosome occupancy than do non-RNAi target transcripts, supporting the view that stalled spliceosomes are a signal for RNAi.

Splicing Intermediates May Be a Preferred Substrate for siRNA Biogenesis

Our finding that inefficient progression of *CNAG_7888* through the splicing pathway increases its targeting by RNAi raised the possibility that splicing intermediates of this transcript, such as the intron 2 lariat intermediate, represent preferred substrates for siRNA biogenesis. We therefore tested whether siRNA production from *CNAG_7888* requires the lariat debranching enzyme (Dbr1), which debranches discarded lariat intermediates and excised lariats in order to allow processing of these RNA species (Chapman and Boeke, 1991; Mayas et al., 2010). We found that Dbr1 was absolutely required for the siRNA production stimulated by a 3′ splice site mutation of *CNAG_7888* intron 2 (Figure 7A). Strikingly, we also observed that debranchase was essential for accumulation of siRNA corresponding to the strong RNAi target transcripts *CNAG_6757*, *CNAG_6844*, and *CNAG_6705* (Figure 7B). These results suggest that debranching of splicing intermediates may enable their use as preferred templates for siRNA production. Such a model predicts that the first exon of either a single- or multi-intron pre-mRNA would

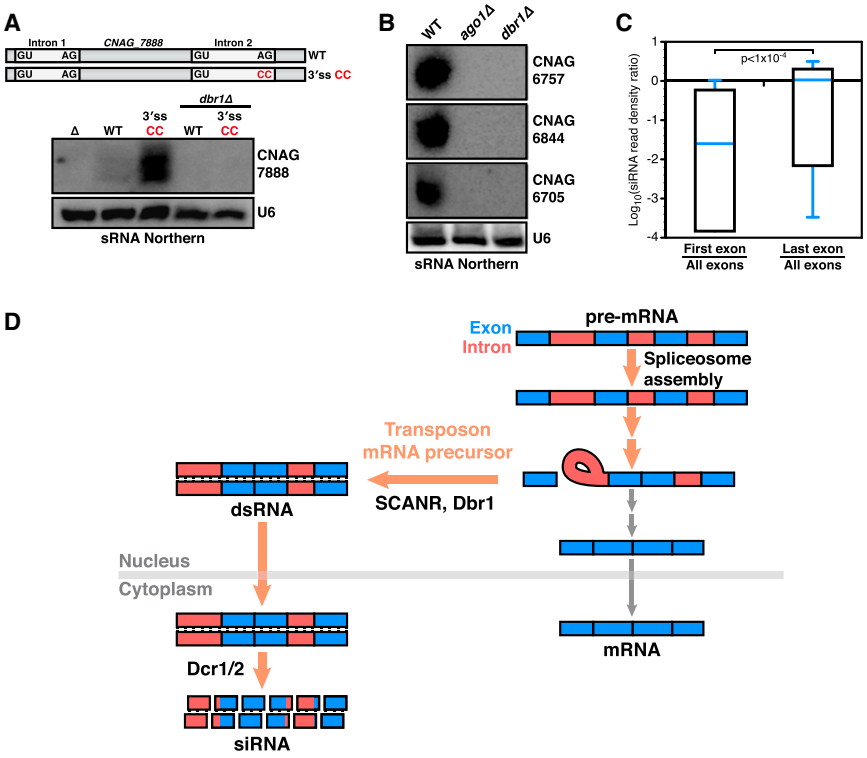


Figure 7. Lariat Debranchase Is Required for siRNA Production

(A) siRNA accumulation triggered by a 3′ splice site mutation of *CNAG_7888* requires the lariat debranching enzyme. RNA was isolated from wild-type or *dbr1Δ* cells that expressed, under the control of a *GAL7* promoter, wild-type *CNAG_7888* or a mutated form of the gene.

(B) sRNA Northern blot. Loading control: U6 snRNA.

(C) Relative siRNA read density in the first or last exon of RNAi target genes. For each of the top 50 RNAi target genes, the following ratio was calculated: siRNA read density in the first (or last) exon over siRNA read density across all exons of the corresponding ORF. These values were $\log_{10}$ transformed and are presented as box plots; blue line indicates median. Values for first versus last exons were compared by the two-tailed Mann-Whitney U test.

(D) Kinetic competition model. The utilization of particular mRNA precursors by SCANR is influenced by their splicing efficiency due to a kinetic competition between splicing and dsRNA synthesis: transcripts that are inefficiently spliced, such as foreign genetic elements, exhibit increased spliceosome association, thereby facilitating their conversion to dsRNA, and, ultimately, siRNA. In this hypothetical example, a lariat intermediate is produced by stalled splicing of a transcript's first intron; downstream introns remain incompletely spliced. This intermediate is acted upon by SCANR and Dbr1 to produce a dsRNA substrate for Dcr1/2. See text for details.

be an underrepresented substrate for siRNA biogenesis because this region cannot be encoded by any lariat intermediate. Consistent with this prediction, we found that, among the top 50 RNAi target transcripts, the siRNA read density in annotated first exons—relative to the overall exonic density in their corresponding genes—was significantly less than the same metric calculated for annotated last exons (Figure 7C). Together, these observations suggest that debranching of splicing intermediates stalled on spliceosomes may be a generally important step in siRNA production.

DISCUSSION

RNAi-related systems that target transposons are important guardians of genome integrity, but our understanding of how self and nonself DNA are distinguished is incomplete. Whereas piRNA systems offer an adaptive mechanism for transposon recognition, the existence of innate mechanisms can also be inferred, particularly for cases in which siRNA is produced against transposons that form dsRNA. Our studies of the human pathogenic yeast *Cryptococcus neoformans* provide multiple lines of evidence that the process of pre-mRNA splicing by the spliceosome plays an important role in defining targets for the RNAi machinery.

Specifically, we find that transcripts targeted by RNAi are stalled in spliceosomes and that stalling promotes siRNA synthesis through a spliceosome-associated RNAi complex, SCANR. Our results are compatible with a competition between the SCANR-mediated recognition of unspliced mRNA precursors and the completion of their splicing, in which inefficiently spliced transcripts are channeled toward dsRNA synthesis. Although previous studies in *S. pombe*, *C. elegans*, and *A. thaliana* have demonstrated that mutation or knockdown of RNA processing factors, including essential splicing factors, can affect siRNA levels, in no case has it been clarified whether these effects are direct, in part because the underlying mechanisms have not been elucidated (Bayne et al., 2008; Herr et al., 2006; Kim et al., 2005; Robert et al., 2005; Tabach et al., 2012). Below, we summarize evidence that intron-containing mRNA precursors are preferred substrates for siRNA synthesis. We discuss how this finding, together with the observation that transposon transcripts display unusually high spliceosome occupancy, suggests a kinetic competition model for RNAi targeting, and we describe potential advantages of such a mechanism. Finally, we speculate on the evolutionary forces that may have led to the high spliceosome occupancy observed for transposon RNAs.

Intron-Containing mRNA Precursors Are a Template for siRNA in *C. neoformans*

Our siRNA sequencing data demonstrate that siRNAs map to the introns and intron-exon junctions of most RNAi target genes, suggesting that either DNA or intron-containing mRNA precursors are a template for siRNA synthesis. Several lines of evidence point to mRNA precursors as relevant templates. First, our studies of SCANR show that this RNA-dependent RNA

polymerase complex is essential for siRNA production and associates with the spliceosome, which itself assembles on pre-mRNA. More importantly, we show that deletion of introns from a strong RNAi target gene reduces the level of its corresponding siRNA, whereas introduction of a 3′ splice site mutation into a weak RNAi target gene, which stalls its splicing, increases the level of its corresponding siRNA. Crucially, the latter effect is blocked by a 5′ splice site mutation that prevents spliceosome association with the transcript. Finally, we find that the lariat debranching enzyme, which acts on branched RNA species generated by the action of the spliceosome (both lariat intermediate and excised lariat), is required for siRNA accumulation. Taken together, these data argue that intron-containing mRNA precursors are templates for dsRNA synthesis and siRNA production.

Spliceosomal Stalling of Transposon Transcripts Suggests a Kinetic Competition

A second key finding of our studies is that RNAi target transcripts display much higher occupancy on spliceosomal complexes than do transcripts that do not enter the RNAi pathway. Importantly, RNAi itself is not responsible for this difference, as cells lacking siRNA display the same behavior. These findings suggest a kinetic competition model for the specification of transposon transcripts as substrates for siRNA synthesis (Figure 7D). In this model, mRNA precursors stalled on the spliceosome remain available for recognition by SCANR, whereas transcripts that complete splicing undergo spliceosome disassembly, intron degradation, and export out of the nucleus. The nuclear localization of SCANR and its physical association with the spliceosome would facilitate the proposed competition. Such a spliceosome-based mechanism is advantageous in that it does not respond only to particular sequences but rather to splicing efficiency itself and, therefore, could in principle recognize multiple types of suboptimal introns, including those of foreign genetic elements not previously encountered.

Although our experiments addressing transcript occupancy on Prp19-containing spliceosomal complexes do not speak to what specific kinetic steps are delayed during the splicing of transposon transcripts, other results suggest that spliceosome stalling events subsequent to the first catalytic step may be particularly capable of engaging SCANR. In fact, we find that a 3′ splice site mutation of *CNAG_7888*, which stalls splicing at the second catalytic step, promotes siRNA production from this locus. Furthermore, this siRNA production requires the lariat debranching enzyme, as does the production of siRNA corresponding to several other endogenous RNAi targets. Because loss of Dbr1 in other systems causes an accumulation of lariat RNAs but does not perturb splicing or spliceosome disassembly (Mayas et al., 2010; Nam et al., 1997), this finding suggests that lariat RNAs generated by the first step of splicing, once debranched, may be preferred substrates for Rdp1 and/or Dicer in *C. neoformans*.

The Inefficient Splicing of Pre-mRNAs that Contain Transposons

The genome of *C. neoformans* is rich in introns, as more than 97% of annotated genes contain at least one intron. In other organisms with intron-rich genomes, such as mammals, splicing enhances gene expression by promoting 3′ end formation, nuclear export of mRNA, and mRNA translation (Lu and Cullen, 2003; Valencia et al., 2008). Intron-dependent gene expression is also a feature of basidiomycetous yeast, raising the possibility that successful transposons in *C. neoformans* require introns (Burns et al., 2005; Lugones et al., 1999). In fact, all of the 12 genes most targeted by siRNA in *C. neoformans*, the majority of which are related to transposons, contain introns. Yet all of these transcripts exhibit abnormally high spliceosome occupancy, suggesting that their introns may be poorly spliced. Furthermore, RNAi target transcripts generally exhibit sequence features predictive of inefficient splicing, including relatively low 5′ splice site strength and increased intron length. But why would these foreign genetic elements tend to be poorly spliced as compared to endogenous genes? We suggest two nonmutually exclusive possibilities.

First, transposons that have entered the *C. neoformans* genome recently by horizontal transfer have had limited time to adapt to its specific splicing preferences. Efficient splicing in *C. neoformans* likely requires appropriate intron size, consensus-matched splice site sequences, and proper exonic splicing enhancer sequences (Warnecke et al., 2008); these particular splicing preferences differ among organisms, creating a barrier to the efficient expression of horizontally transferred genes. Thus, the suboptimal 5′ splice sites and intron sizes of *C. neoformans* RNAi targets could be due in part to their limited coevolution with endogenous genes.

A second possibility is based on an extensive literature describing cryptic introns in transposons that appear to minimize the impact of transposon insertion into host genes (Purugganan, 2002). Transposon disruptions of essential host genes can result in the host cell's death, which is deleterious for both the transposon and the host. These same transposon insertions, however, can be viable if transposon sequences are spliced out of the essential gene's transcript. Therefore, splicing allows transposons to circumvent some negative fitness consequences associated with their propagation (Gierl, 1990). Consistent with this idea, multiple transposons—including *Ds* in maize, the 412 retrotransposon and P elements in *Drosophila*, and *Tc1* in *C. elegans*—contain splice sites near their termini such that the whole transposon can be spliced out of a larger transcription unit, thereby limiting disruption of endogenous loci (Belancio et al., 2008; Purugganan and Wessler, 1992; Purugganan, 2002). Many other transposons, including Harbinger family DNA transposons and LINE elements, also contain introns, although these introns do not span the entire transposon (Belancio et al., 2006; Kapitonov and Jurka, 2004). If a transposon, after its insertion into a host gene, is spliced in such a way that mitigates its negative effects on host gene expression but also removes sequences important for mobilization, then weak splicing would be favored as a compromise between transposon expression and host organism health (Menssen et al., 1990). Such inefficiently spliced transcripts may become stalled in the spliceosome and recognized by SCANR, thereby contributing to the high spliceosome occupancy of transposon transcripts and to the production of transposon-specific siRNA for genome defense. Whether these principles apply to other organisms

remains to be elucidated, but we are intrigued by the recent report of a class of *C. elegans* endo-siRNA that appears to be derived from intron-containing mRNA precursors (Warf et al., 2012).

EXPERIMENTAL PROCEDURES

Yeast Strains

Yeast strains used in this study are listed in Table S4. All *C. neoformans* strains were derived from strain H99 using standard procedures (Chun and Madhani, 2010).

Tandem Affinity Protein Purification

C. neoformans cultures grown in YPAD media were harvested, snap frozen, and then lysed using a coffee grinder and mortar and pestle. Proteins tagged with CBP-2xFLAG were purified using anti-FLAG M2 resin (Sigma) and calmodulin resin (Stratagene) according to manufacturers' instructions. Purified proteins were analyzed by immunoblot or mass spectrometry as described in Extended Experimental Procedures and Table S6.

RNA Immunoprecipitation

Epitope-tagged proteins were purified using anti-FLAG resin as described above, and associated RNA was isolated by phenol-chloroform extraction, as described in Extended Experimental Procedures. Table S5 lists primers used for RT-qPCR.

RNA Isolation and sRNA Northern Blot

Total RNA was isolated using TRIzol (Invitrogen), whereas small RNA was isolated using a modified mirVana (Ambion) protocol, as described in Extended Experimental Procedures. Small RNA Northern blots were performed as described previously (Pall and Hamilton, 2008).

ACCESSION NUMBERS

Small RNA sequencing data were deposited in the Gene Expression Omnibus under accession number GSE43363.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, one data file, and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.01.046.

ACKNOWLEDGMENTS

We thank members of the Madhani and Guthrie labs, P. O'Farrell, J. Reuter, and S. Coyle for critical reading of the manuscript and helpful discussions; N. Nguyen for media preparation; and X. Wang and J. Heitman for strains and plasmids. Mass spectrometry experiments were supported by the National Center for Research Resources (5P41RR011823-17), the National Institute of General Medical Sciences (8 P41 GM103533-17), and the National Institute on Aging (R01AG027463-04). P.A.D. and H.D.M. designed the study. C.C., I.A.D., and D.P.B. performed high-throughput sequencing shown in Figure 1. J.T., J.J.M., and J.R.Y. performed mass spectrometry analyses shown in Figures 2, 3, and 4. P.N. performed subcellular localization of tagged proteins shown in Figures 2, 3, 4, and S1. B.J.S. performed bioinformatic analyses shown in Figures 1, 5, and 7 and Tables S1 and S2. P.A.D. performed experiments shown in Figures 2, 3, 4, 5, 6, 7, S2, S3, and S4. P.A.D. and H.D.M. wrote the manuscript. All authors contributed to editing the manuscript.

Received: July 11, 2012
Revised: November 13, 2012
Accepted: January 17, 2013
Published: February 14, 2013

REFERENCES

Bayne, E.H., Portoso, M., Kagansky, A., Kos-Braun, I.C., Urano, T., Ekwall, K., Alves, F., Rappsilber, J., and Allshire, R.C. (2008). Splicing factors facilitate RNAi-directed silencing in fission yeast. Science *322*, 602–606.

Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. Cold Spring Harb. Symp. Quant. Biol. *71*, 523–530.

Belancio, V.P., Hedges, D.J., and Deininger, P. (2006). LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic Acids Res. *34*, 1512–1521.

Belancio, V.P., Roy-Engel, A.M., and Deininger, P. (2008). The impact of multiple splice sites in human L1 elements. Gene *411*, 38–45.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089–1103.

Burns, C., Gregory, K.E., Kirby, M., Cheung, M.K., Riquelme, M., Elliott, T.J., Challen, M.P., Bailey, A., and Foster, G.D. (2005). Efficient GFP expression in the mushrooms Agaricus bisporus and Coprinus cinereus requires introns. Fungal Genet. Biol. *42*, 191–199.

Cam, H.P., Chen, E.S., and Grewal, S.I. (2009). Transcriptional scaffolds for heterochromatin assembly. Cell *136*, 610–614.

Cellini, A., Felder, E., and Rossi, J.J. (1986). Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3′ splice site. EMBO J. *5*, 1023–1030.

Chapman, K.B., and Boeke, J.D. (1991). Isolation and characterization of the gene encoding yeast debranching enzyme. Cell *65*, 483–492.

Chua, K., and Reed, R. (2001). An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing. Mol. Cell. Biol. *21*, 1509–1514.

Chun, C.D., and Madhani, H.D. (2010). Applying genetics and molecular biology to the study of the human pathogen Cryptococcus neoformans. Methods Enzymol. *470*, 797–831.

Cogoni, C., and Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature *399*, 166–169.

Conley, A.B., Miller, W.J., and Jordan, I.K. (2008). Human cis natural antisense transcripts initiated by transposable elements. Trends Genet. *24*, 53–56.

Connelly, C., and Hieter, P. (1996). Budding yeast SKP1 encodes an evolutionarily conserved kinetochore protein required for cell cycle progression. Cell *86*, 275–285.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science *326*, 544–550.

El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., Hakimi, M.A., Jacobsen, S.E., Cooke, R., and Lagrange, T. (2007). Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev. *21*, 2539–2544.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. *10*, 94–108.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., and Zamore, P.D. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077–1081.

Gierl, A. (1990). How maize transposable elements escape negative selection. Trends Genet. *6*, 155–158.

Herr, A.J., Molnàr, A., Jones, A., and Baulcombe, D.C. (2006). Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. Proc. Natl. Acad. Sci. USA *103*, 14994–15001.

Hogg, R., McGrail, J.C., and O'Keefe, R.T. (2010). The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. Biochem. Soc. Trans. *38*, 1110–1115.

Hughes, S.S., Buckley, C.O., and Neafsey, D.E. (2008). Complex selection on intron size in Cryptococcus neoformans. Mol. Biol. Evol. *25*, 247–253.

Irimia, M., Penny, D., and Roy, S.W. (2007). Coevolution of genomic intron number and splice sites. Trends Genet. *23*, 321–325.

Janbon, G., Maeng, S., Yang, D.H., Ko, Y.J., Jung, K.W., Moyrand, F., Floyd, A., Heitman, J., and Bahn, Y.S. (2010). Characterizing the role of RNA silencing components in Cryptococcus neoformans. Fungal Genet. Biol. *47*, 1070–1080.

Kapitonov, V.V., and Jurka, J. (2004). Harbinger transposons and an ancient HARBI1 gene derived from a transposase. DNA Cell Biol. *23*, 311–324.

Kelly, W.G., and Aramayo, R. (2007). Meiotic silencing and the epigenetics of sex. Chromosome Res. *15*, 633–651.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133–141.

Khurana, J.S., Wang, J., Xu, J., Koppetsch, B.S., Thomson, T.C., Nowosielska, A., Li, C., Zamore, P.D., Weng, Z., and Theurkauf, W.E. (2011). Adaptation to P element transposon invasion in Drosophila melanogaster. Cell *147*, 1551–1563.

Kim, J.K., Gabel, H.W., Kamath, R.S., Tewari, M., Pasquinelli, A., Rual, J.F., Kennedy, S., Dybbs, M., Bertin, N., Kaplan, J.M., et al. (2005). Functional genomic analysis of RNA interference in C. elegans. Science *308*, 1164–1167.

Lee, H.C., Aalto, A.P., Yang, Q., Chang, S.S., Huang, G., Fisher, D., Cha, J., Poranen, M.M., Bamford, D.H., and Liu, Y. (2010). The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. PLoS Biol. *8*, e1000496.

Loftus, B.J., Fung, E., Roncaglia, P., Rowley, D., Amedeo, P., Bruno, D., Vamathevan, J., Miranda, M., Anderson, I.J., Fraser, J.A., et al. (2005). The genome of the basidiomycetous yeast and human pathogen Cryptococcus neoformans. Science *307*, 1321–1324.

Lu, S., and Cullen, B.R. (2003). Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. RNA *9*, 618–630.

Lugones, L.G., Scholtmeijer, K., Klootwijk, R., and Wessels, J.G. (1999). Introns are necessary for mRNA accumulation in Schizophyllum commune. Mol. Microbiol. *32*, 681–689.

Maiti, M., Lee, H.C., and Liu, Y. (2007). QIP, a putative exonuclease, interacts with the Neurospora Argonaute protein and facilitates conversion of duplex siRNA into single strands. Genes Dev. *21*, 590–600.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell *136*, 656–668.

Mayas, R.M., Maita, H., Semlow, D.R., and Staley, J.P. (2010). Spliceosome discards intermediates via the DEAH box ATPase Prp43p. Proc. Natl. Acad. Sci. USA *107*, 10020–10025.

Menssen, A., Höhmann, S., Martin, W., Schnable, P.S., Peterson, P.A., Saedler, H., and Gierl, A. (1990). The En/Spm transposable element of Zea mays contains splice sites at the termini generating a novel intron from a dSpm element in the A2 gene. EMBO J. *9*, 3051–3057.

Nam, K., Lee, G., Trambley, J., Devine, S.E., and Boeke, J.D. (1997). Severe growth defect in a Schizosaccharomyces pombe mutant defective in intron lariat degradation. Mol. Cell. Biol. *17*, 809–818.

Nolan, T., Cecere, G., Mancone, C., Alonzi, T., Tripodi, M., Catalanotto, C., and Cogoni, C. (2008). The RNA-dependent RNA polymerase essential for post-transcriptional gene silencing in Neurospora crassa interacts with replication protein A. Nucleic Acids Res. *36*, 532–538.

Pall, G.S., and Hamilton, A.J. (2008). Improved northern blot method for enhanced detection of small RNA. Nat. Protoc. *3*, 1077–1084.

Purugganan, M.D. (2002). The splicing of transposable elements: evolution of a nuclear defense against genomic invaders? In Horizontal Gene Transfer, M. Syvanen and C.I. Kado, eds. (San Diego, CA: Academic Press), pp. 187–195.

Purugganan, M., and Wessler, S. (1992). The splicing of transposable elements and its role in intron evolution. Genetica *86*, 295–303.

Robert, V.J., Sijen, T., van Wolfswinkel, J., and Plasterk, R.H. (2005). Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. Genes Dev. *19*, 782–787.

Shabalina, S.A., and Koonin, E.V. (2008). Origins and evolution of eukaryotic RNA interference. Trends Ecol. Evol. *23*, 578–587.

She, X., Xu, X., Fedotov, A., Kelly, W.G., and Maine, E.M. (2009). Regulation of heterochromatin assembly on unpaired chromosomes during Caenorhabditis elegans meiosis by components of a small RNA-mediated pathway. PLoS Genet. *5*, e1000624.

Shiu, P.K., Raju, N.B., Zickler, D., and Metzenberg, R.L. (2001). Meiotic silencing by unpaired DNA. Cell *107*, 905–916.

Sijen, T., and Plasterk, R.H. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature *426*, 310–314.

Tabach, Y., Billi, A.C., Hayes, G.D., Newman, M.A., Zuk, O., Gabel, H., Kamath, R., Yacoby, K., Chapman, B., Garcia, S.M., et al. (2012). Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature. Published online December 23, 2012. http://dx.doi.org/10.1038/nature11779.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123–132.

Valencia, P., Dias, A.P., and Reed, R. (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc. Natl. Acad. Sci. USA *105*, 3386–3391.

Wang, X., Hsueh, Y.P., Li, W., Floyd, A., Skalsky, R., and Heitman, J. (2010). Sex-induced silencing defends the genome of Cryptococcus neoformans via RNAi. Genes Dev. *24*, 2566–2582.

Warf, M.B., Shepherd, B.A., Johnson, W.E., and Bass, B.L. (2012). Effects of ADARs on small RNA processing pathways in C. elegans. Genome Res. *22*, 1488–1498.

Warnecke, T., Parmley, J.L., and Hurst, L.D. (2008). Finding exonic islands in a sea of non-coding sequence: splicing related constraints on protein composition and evolution are common in intron-rich genomes. Genome Biol. *9*, R29.

Yang, N., and Kazazian, H.H., Jr. (2006). L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat. Struct. Mol. Biol. *13*, 763–771.

EXTENDED EXPERIMENTAL PROCEDURES

Yeast Media and Techniques

Strains were grown in YPAD (1% yeast extract, 2% Bacto-peptone, 2% glucose, 0.015% L-tryptophan, 0.004% adenine), YPAG (1% yeast extract, 2% Bacto-peptone, 2% galactose, 0.015% L-tryptophan, 0.004% adenine), or YNB (1.5 g/l yeast nitrogen base, 5 g/l ammonium sulfate, 2% glucose) media at 30°C. Because *C. neoformans* can respond to light, strains were grown and harvested in darkness (Idnurm and Heitman, 2005).

Gene Nomenclature

C. neoformans genes were identified using Broad Institute (Cambridge, MA) annotations of the *var. grubii* H99 sequence (http://www.broadinstitute.org/annotation/genome/cryptococcus_neoformans/MultiHome.html), in which genes are named "CNAG_#."

Small RNA Library Preparation

cDNA libraries were prepared from small RNAs as described (Grimson et al., 2008) and sequenced using the Illumina SBS platform.

siRNA Read Processing

Sequencing reads that passed a quality filter were truncated at the 3′ linker sequence (TCGTAT) and then mapped to the loci encoding rRNA and tRNA genes, allowing up to one mismatch and randomly sampling multiple alignments where applicable. Sequences that did not align to the rRNA or tRNA were then aligned against the full genome, allowing only perfect matches and randomly sampling multiple alignments where applicable. Sequence and feature files for *C. neoformans var. grubii* H99 were obtained from the Broad Institute (Cambridge, MA) on January 20, 2011. Mapping was done using Bowtie version 0.12.7 (Langmead et al., 2009).

Scripts for siRNA Read Alignment

The three Python scripts that were used to prepare sequencing tagcounts for alignment to the *C. neoformans* genome such that reads mapping to multiple genomic locations could be randomly distributed can be found in Data S1.

siRNA Read Classification

Genomic regions giving rise to siRNAs in wild-type cells were identified as follows. The genome of *C. neoformans* was parsed into non-overlapping 100 bp windows. Windows with high levels of siRNA reads were selected by applying a read density cutoff of $\geq$ 10 reads/window. Adjacent windows passing the cutoff were merged. siRNAs in these windows were then classified based on their genomic positions (Figure 1D). Centromeric sequences, which are known to consist of fragments of transposable elements, were used as queries to identify centromere-like sequence a windows, which by definition align to centromeric sequences with a tblastx E-value cutoff of 0.00001.

Informatic Analysis of Intron Features and siRNA Read Density

Predicted intron lengths and 5′ splice site sequences were obtained from the current Broad annotation of the *C. neoformans var grubii* genome sequence. 5′ splice site sequences were converted into self-information (bits). Comparisons of siRNA targets to the genome overall was performed using a two-tailed Kolmogorov–Smirnov test.

Comparison of siRNA read density in first versus last exons was performed as follows. The siRNA read density in the first (or last) exon of each gene was normalized to the siRNA read density in the entire corresponding ORF. The values were then $\log_{10}$-transformed (setting any 0 values equal to the minimum nonzero value for a given exon class) and compared using a two-tailed Mann-Whitney U test.

Tandem Affinity Protein Purification

To purify proteins tagged with CBP-2xFLAG, *C. neoformans* cultures were grown to OD_{600} = 2.0 in YPAD media, at which point they were harvested, resuspended in TAP buffer (25 mM HEPES-KOH pH7.9, 0.1 mM EDTA, 0.5 mM EGTA, 2 mM $MgCl_2$, 20% glycerol, 0.1% Tween-20, 300 mM KCl, 1x EDTA-free Complete protease inhibitor (CPI; Roche)), snap frozen, then lysed using a coffee grinder (3 min) and mortar and pestle (20 min). The frozen powder was resuspended in TAP buffer and cleared by 27,000 x *g* centrifugation for 40 min at 4°C. Anti-FLAG M2 affinity resin (Sigma) was incubated in cleared lysate for 2 hr at 4°C, at which point the resin was washed three times with TAP buffer. Tagged protein was eluted by three washes with FLAG elution buffer at 4°C (25 mM HEPES-KOH pH7.9, 2 mM $MgCl_2$, 20% glycerol, 300 mM KCl, 1x CPI, 0.4 mg/ml 3xFLAG peptide (Sigma)) totaling 1 hr. For the second purification step, 5 volumes of calmodulin binding buffer (10 mM Tris-HCl pH7.9, 10 mM β-mercaptoethanol, 2 mM $CaCl_2$, 0.1% Triton X-100, 300 mM NaCl, 1x CPI) were added to the anti-FLAG resin eluate; the resulting solution was incubated with calmodulin beads (Stratagene) at 4°C overnight. The beads were then washed once with calmodulin binding buffer and three times with calmodulin wash buffer (same as calmodulin binding buffer, except 0.1 mM $CaCl_2$) at 4°C. Protein was eluted by five washes with calmodulin elution buffer at 4°C (10 mM Tris-HCl, 10 mM β-mercaptoethanol, 3 mM EGTA, 0.1% Triton X-100, 300 mM NaCl) totaling 1 hr 45 min. Eluted protein was precipitated with 13% trichloroacetic acid, washed with acetone, and analyzed by mass spectrometry or immunoblot.

Mass Spectrometry Reagents and Chemicals

Unless otherwise noted all chemicals were purchased from Thermo Fisher Scientific. Deionized water (18.2 MW, Barnstead) was used for all preparations. Buffer A was 5% acetonitrile 0.1% formic acid, B was 80% acetonitrile 0.1% formic acid, and C was 500 mM ammonium acetate.

Protein Digestion

Proteins were reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride (C4706, Sigma) and alkylated with 10 mM Iodoacetamide (Sigma). Proteins were digested for 18 hr at 37°C in 2 M urea, 100 mM Tris pH 8.5, 1 mM $CaCl_2$ with 1 μg trypsin (Promega). Digest was stopped with formic acid, 5% final concentration. Debris was removed by centrifugation, 30 min 18,000 x *g*.

MudPIT Microcolumn

A MudPIT microcolumn (Washburn et al., 2001; Wolters et al., 2001) was prepared by first creating a Kasil frit at one end of an undeactivated 250 mm ID/360 mm OD capillary (Agilent Technologies, Inc.). The Kasil frit was prepared by briefly dipping a 20-30 cm capillary in well-mixed 300 ml Kasil 1624 (PQ Corporation) and 100 ml formamide, curing at 100°C for 4 hr, and cutting the frit to ~2 mm in length. Strong cation exchange particles (SCX Luna, 5 mm dia., 125 Å pores, Phenomenex) were packed in-house from particle slurries in methanol to 2.5 cm. 2 cm reversed phase particles (C18 Aqua, 3 mm dia., 125 Å pores, Phenomenex) were then successively packed onto the capillary using the same method as SCX loading.

MudPIT Analysis

An analytical RPLC column was generated by pulling a 100 mm ID/360 mm OD capillary (Polymicro Technologies) to 5 mm ID tip. Reversed phase particles (Luna C18, 3 mm dia., 125 Å pores, Phenomenex) were packed directly into the pulled column at 800 psi until 15 cm long. The column was further packed, washed, and equilibrated at 100 bar with buffer B followed by buffer A. MudPIT and analytical columns were assembled using a zero-dead volume union (Upchurch Scientific). LC-MS/MS analysis was performed using an Agilent 1100 HPLC pump and Finnigan LTQ using an in-house built electrospray stage. Electrospray was performed directly from the analytical column by applying the ESI voltage at a tee (150 mm ID, Upchurch Scientific) directly downstream of a 1:1000 split flow used to reduce the flow rate to 300 nl/min through the columns. 5-step MudPIT experiments were performed where each step corresponds to 0, 20, 50, 80, and 100% buffer C being run for 5 min at the beginning of a 110 min gradient. Precursor scanning was performed from 300 - 2000 m/z. Data-dependent acquisition of MS/MS spectra was performed with the following settings: MS/MS on the 5 most intense ions per precursor scan. Dynamic exclusion settings used were as follows: repeat count, 1; repeat duration, 30 s; exclusion list size, 300; and exclusion duration, 180 s.

Protein and peptide identification and modified peptide analysis were done with Integrated Proteomics Pipeline - IP2 (Integrated Proteomics Applications, Inc.) using ProLuCID, DTASelect2. Spectrum raw files were extracted into ms2 files from raw files using RawExtract 1.9.9 (http://fields.scripps.edu/downloads.php) (McDonald et al., 2004), and the tandem mass spectra were searched against a *Cryptococcus* proteins database (Broad Institute, Cambridge, MA). In order to accurately estimate peptide probabilities and false discovery rates, we used a decoy database containing the reversed sequences of all the proteins appended to the target database (Peng et al., 2003). Tandem mass spectra were matched to sequences using the ProLuCID algorithm with 600 ppm peptide mass tolerance. ProLuCID searches were done on an Intel Xeon cluster running under the Linux operating system. The search space included all fully tryptic peptide candidates that fell within the mass tolerance window with no miscleavage constraint. Carbamidomethylation (+57.02146 Da) of cysteine was considered as a static modification. DTASelect parameters were -p 2 -y 0–trypstat–dm -in.

Background Filtering Criteria for Mass Spectrometry Analysis

To remove likely contaminants from the list of proteins identified by mass spectrometry, the data set was filtered to remove proteins that were: (1) identified by less than 10% peptide coverage, (2) structural components of the ribosome, (3) proteins identified in untagged sample, or (4) other likely-abundant proteins such as cytoskeletal proteins, metabolic proteins, chaperones, and mitochondrial proteins. Filtered proteins are listed in Table S6.

Immunoblotting

Proteins were analyzed by SDS-PAGE and immunoblotting using primary antibodies at the following concentrations: mouse monoclonal anti-FLAG (Sigma F3165, 1:3,000), rabbit polyclonal anti-Myc (Abcam, 1:2,500), and rabbit polyclonal anti-PSTAIRE (Santa Cruz Biotechnology sc-53, 1:2,500). Secondary antibodies included HRP-conjugated goat anti-mouse (Bio-Rad, 1:20,000) and HRP-conjugated goat anti-rabbit (Bio-Rad, 1:20,000).

RNA Immunoprecipitation

To isolate RNA associated with CBP-2xFLAG-tagged proteins, *C. neoformans* cultures were grown to OD_{600} = 2.0 in YPAD or YPAG media, at which point they were harvested, resuspended in TAP buffer (supplemented with 100 U/ml RNase inhibitor (New England Biolabs)), snap frozen, then lysed using a coffee grinder (3 min) and mortar and pestle (15 min). Upon thawing, the lysate was cleared by centrifugation at 27,000 x *g* for 40 min at 4°C, after which it was incubated with anti-FLAG M2 affinity resin (Sigma) for 2 hr at 4°C.

The resin was washed three times with TAP buffer, then bound proteins were eluted by three washes with FLAG elution buffer at 4°C, totaling 1 hr. To purify protein-associated RNA, the eluate was incubated with 0.17 mg/ml Proteinase K (Sigma) for 25 min at 37°C followed by an acid phenol-chloroform extraction and ethanol precipitation. RNA samples were treated with DNaseI (DNA-free, Ambion), and subsequent RT-qPCR analysis was carried out using primers listed in Table S5.

RNA Isolation and sRNA Northern Blot

To analyze gene expression or siRNA abundance, *C. neoformans* cultures were grown to OD_{600} = 1.0 in YPAD or YPAG media, at which point they were harvested and snap frozen. RNA was isolated using TRIzol (Invitrogen). To measure transcript abundance, total RNA was treated with DNaseI (Roche) followed by RT-qPCR or primer extension. To measure siRNA abundance, small RNAs were first enriched from total RNA samples by performing a modified mirVana (Ambion) small RNA isolation procedure, as described previously (Gu et al., 2011). Next, 40 μg sRNA samples were resolved in a 15% polyacrylamide gel and transferred to a Hybond-NX membrane (Amersham), which was incubated in crosslinking solution (0.16 M *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (Sigma) prepared in 0.13 M 1-methylimidazole at pH8) at 60°C for 1 hr (Pall and Hamilton, 2008). The crosslinked membrane was washed with water and blocked with Ultrahyb solution (Ambion) for 30 min at 68°C. Radiolabeled riboprobes corresponding to the sense strand of particular gene loci were generated by in vitro transcription of a linearized plasmid (MAXIscript, Ambion). These riboprobes corresponded to the entire locus of each examined gene except for *CNAG_6705*, in which case only the non-repetitive regions were included. Riboprobes were fragmented by base hydrolysis to an average size of 100 nt, and hybridized to the membrane overnight at 48°C. The membrane was washed 5 min two times at 48°C (2x SSC, 0.1% SDS), then 15 min two times at 48°C (0.1x SSC, 0.1% SDS) and imaged using a storage phosphor screen (Amersham).

RT-qPCR

cDNA was generated by reverse transcription of 10 μg DNaseI-treated, total RNA by SuperScript III reverse transcriptase (Invitrogen) using oligo-$dT_{20}N$ (38 ng/μl) and random 9-mers (10 ng/μl) as primers. The manufacturer's standard reactions conditions were used.

Primer Extension

For primer extension, ^{32}P end-labeled primers were annealed to 15 μg total RNA and extended by AMV reverse transcriptase for 1 hr at 42°C (Primer Extension System, Promega), after which RNA was eliminated by base hydrolysis. Products were resolved in a 6% denaturing polyacrylamide gel and sized relative to a ΦX174 DNA/Hinf I ladder (Promega).

Fluorescence Microscopy

Strains expressing mCherry and GFP fusion proteins were grown to saturation in YPAD (or YPAG) overnight, spotted on V8 mating medium (or V8 with 2% galactose) and incubated at 25°C for 48 hr. Cells were scraped off the plates, resuspended in water, and imaged immediately at 63x magnification in an Axiovert 200 M (Zeiss) microscope running Axiovision software. The images were pseudocolored and cropped using Photoshop software (Adobe). Subcellular localization was quantified by assessing cytoplasmic and nuclear signal in 100 cells per genotype. For nuclear staining, cells were fixed using 2% paraformaldehyde for 10 min at room temperature. After washing with 1x PBS, nuclei were stained using Hoechst 33342 (Invitrogen) at a concentration of 10 μg/ml. Cells were washed and imaged as described above. In order to visualize P-bodies, log phase cells grown in YNB with 2% galactose were spun down, washed, and incubated in YNB without galactose for 10 min. Cells were washed again, resuspended in media with galactose and imaged immediately.

mCherry-Ago1, Gwo1-mCherry, and Srr1-mCherry were expressed from their endogenous promoters, whereas GFP-Gwc1, GFP-Qip1, GFP-Dcp1, and GFP-Gwo1 were expressed from the *GAL7* promoter to facilitate detection.

Yeast Two-Hybrid Analysis

Plasmids were generated that encoded *C. neoformans* proteins (Rdp1, Ago1, Gwc1, Qip1, or Gwo1) fused to either a transcriptional activation domain (AD) or a LexA DNA-binding domain. Plasmids encoding AD fusion proteins, which were expressed from the *GAL1* promoter, were used to transform the *S. cerevisiae* strain W303A, whereas plasmids encoding LexA fusion proteins, which were expressed from the *ADH* promoter, were used to transform EGY48. The EGY48 strain also carried the plasmid pSH18-34, which encodes 8 LexA operator sequences upstream of *LacZ*. To assess *LacZ* expression stimulated by fusion protein interaction, W303A- and EGY48-derived strains were mated and maintained as diploids. Saturated cultures in SC -his -trp -ura +2% raffinose media were diluted 1:20 in SC -his -trp -ura +1% raffinose +2% galactose media and incubated at 30°C with shaking until they reached log phase, at which point β-galactosidase activity was assessed in an Infinite M200 96-well plate reader (Tecan) as previously described (Shock et al., 2009). The β-galactosidase activity stimulated by any given interaction between an AD fusion protein and a LexA fusion protein was normalized to a bait-only control in which the LexA fusion protein was expressed alongside an AD that was not fused to any additional sequence. Interactions that stimulated β-galactosidase activity at least 3-fold relative to the activity of a bait-only control were deemed positive.

SUPPLEMENTAL REFERENCES

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B.J., Chiang, H.R., King, N., Degnan, B.M., Rokhsar, D.S., and Bartel, D.P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature *455*, 1193–1197.

Gu, W., Claycomb, J.M., Batista, P.J., Mello, C.C., and Conte, D. (2011). Cloning Argonaute-associated small RNAs from Caenorhabditis elegans. Methods Mol. Biol. *725*, 251–280.

Idnurm, A., and Heitman, J. (2005). Light controls growth and development via a conserved pathway in the fungal kingdom. PLoS Biol. *3*, e95.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

McDonald, W.H., Tabb, D.L., Sadygov, R.G., MacCoss, M.J., Venable, J., Graumann, J., Johnson, J.R., Cociorva, D., and Yates, J.R., 3rd. (2004). MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. Rapid Commun. Mass Spectrom. *18*, 2162–2168.

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J. Proteome Res. *2*, 43–50.

Shock, T.R., Thompson, J., Yates, J.R., 3rd, and Madhani, H.D. (2009). Hog1 mitogen-activated protein kinase (MAPK) interrupts signal transduction between the Kss1 MAPK and the Tec1 transcription factor to maintain pathway specificity. Eukaryot. Cell *8*, 606–616.

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. *19*, 242–247.

Wolters, D.A., Washburn, M.P., and Yates, J.R., 3rd. (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal. Chem. *73*, 5683–5690.

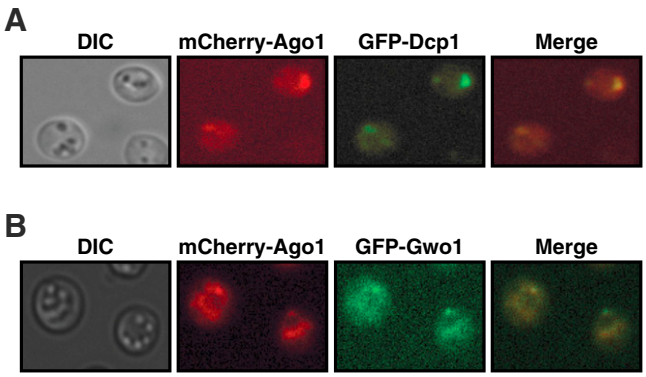


Figure S1. Characterization of Ago1 Cytoplasmic Foci, Related to Figures 2 and 3

(A) Co-localization of Ago1 and Dcp1. mCherry-Ago1 was expressed from its endogenous promoter, whereas GFP-Dcp1, which localizes to P-bodies, was expressed from a *GAL7* promoter to facilitate detection. Unfixed cells expressing both fusion proteins were examined after incubation in YNB media.

(B) Co-localization of Ago1 and Gwo1. mCherry-Ago1 was expressed from its endogenous promoter, whereas GFP-Gwo1 was expressed from the *GAL7* promoter to facilitate detection. Unfixed cells expressing both fusion proteins were examined after incubation in YNB media.

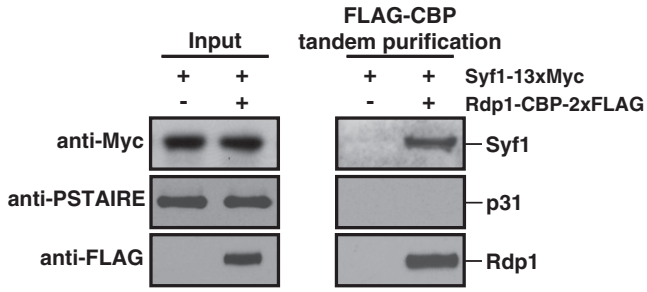


Figure S2. Coimmunoprecipitation of SCANR Subunit Rdp1 and Spliceosome Component Syf1, Related to Figure 4

Coimmunoprecipitation of spliceosome component Syf1 with Rdp1. Strains expressing Syf1-13xMyc and Rdp1-CBP-2xFLAG were subjected to tandem affinity purification using anti-FLAG and calmodulin resins. Input and purified material were analyzed by immunoblot using anti-FLAG, anti-Myc, or anti-PSTAIRE antibody, which stains the negative control protein p31.

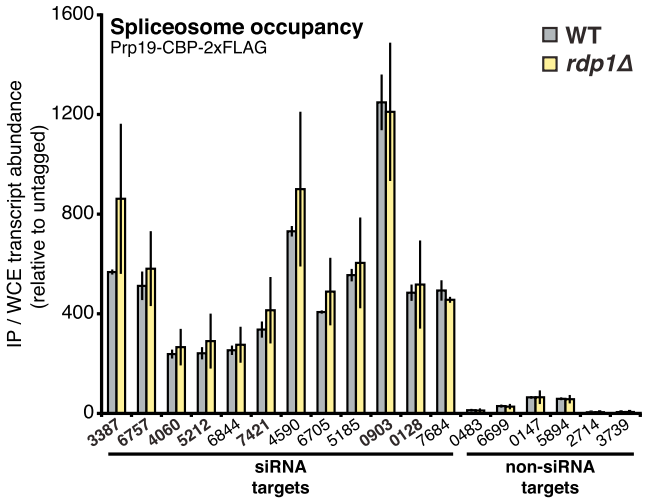


Figure S3. Spliceosome Occupancy of RNAi Target Transcripts in the Absence of siRNA, Related to Figure 5

Association of siRNA target and non-siRNA target transcripts with the spliceosome in the absence of RNAi. Spliceosomes were purified from wild-type or *rdp1Δ* cells by immunoprecipitation of Prp19-CBP-2xFLAG and co-purified RNAs were detected by RT-qPCR. Levels of individual RNAs co-immunoprecipitated with Prp19 were normalized to their abundance in wild-type whole cell extract. IP/WCE values are relative to those of purifications from wild-type (untagged) lysates. Error bars: SD. Bold: transposon-related transcripts.

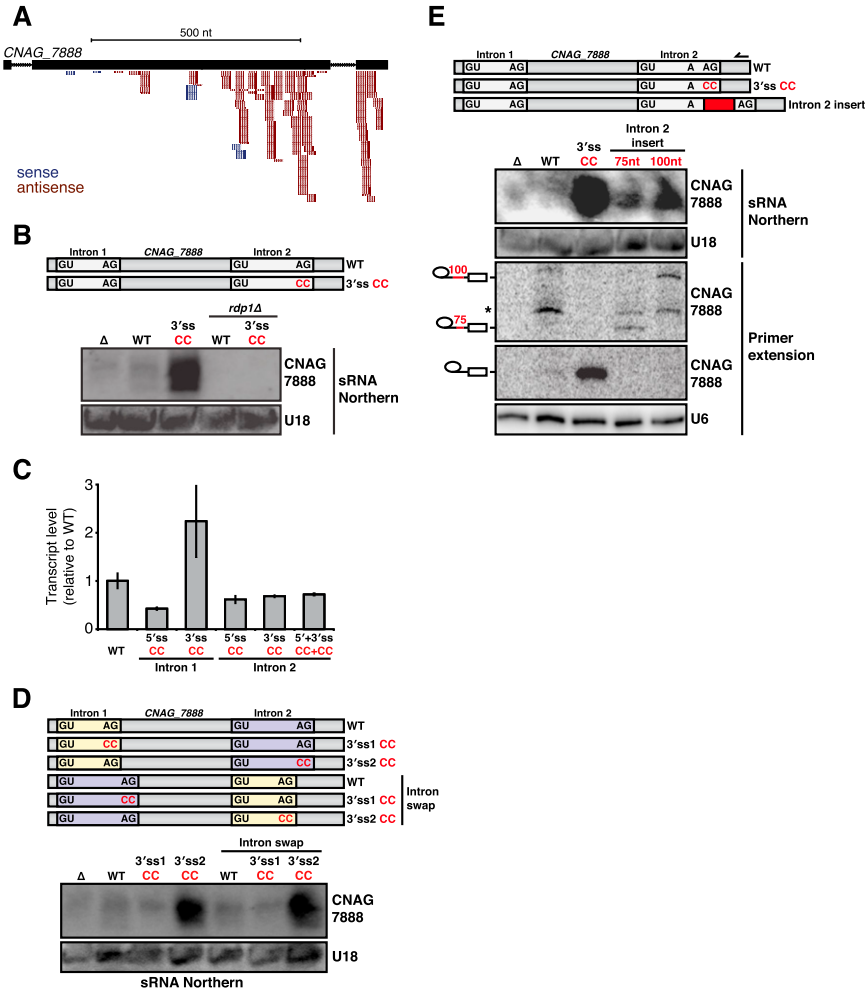


Figure S4. Stalled Spliceosomes Promote siRNA Production by Rdp1, Related to Figure 6

(A) Density plot of siRNAs mapping to the genomic sequence of an RNAi target transcript, *CNAG_7888*, which comprises three exons and two introns.

(B) Rdp1 dependence of siRNA generation triggered by a 3′ splice site mutation of *CNAG_7888* intron 2. RNA was isolated from wild-type or *rdp1Δ* cells that expressed, under the control of a *GAL7* promoter, wild-type *CNAG_7888* or a mutated form of the gene. siRNA derived from *CNAG_7888* was detected by riboprobe hybridization; U18 snoRNA served as loading control.

(C) Transcript levels of wild-type *CNAG_7888* as well as mutated forms of the gene in which individual splice sites were mutated, as assessed by RT-qPCR. All *CNAG_7888* alleles were expressed from a *GAL7* promoter at the endogenous *CNAG_7888* locus. Expression levels were normalized to levels of actin transcript. Error bars: SD.

(D) Effect of intron sequence on the siRNA production triggered by splice site mutations of *CNAG_7888*. RNA was isolated from cells that expressed, from a *GAL7* promoter, either wild-type *CNAG_7888* or a mutated form of the gene in which the sequences of its two introns were swapped and splice sites were mutated. siRNA derived from *CNAG_7888* was detected by riboprobe hybridization; staining of U18 snoRNA served as loading control.

(E) Effect of suboptimal intron size on siRNA production from and splicing of the *CNAG_7888* transcript. RNA was isolated from cells that expressed, from a *GAL7* promoter, either wild-type *CNAG_7888* or a mutated form of the gene containing a sequence insertion between the branchpoint adenine and 3′ splice site of intron 2. siRNA production was assessed by riboprobe hybridization and splicing was assessed by primer extension using a labeled primer complementary to *CNAG_7888* exon 3, which yielded a discrete product corresponding to the lariat intermediate of each mutant intron. Asterisk denotes a nonspecific primer extension product. Primer extension using a primer specific to U6 snRNA served as a loading control.