

The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting

Dongming Liang,^{1,6} Deirdre C. Tatomer,^{1,6} Zheng Luo,² Huang Wu,³ Li Yang,^{2,4} Ling-Ling Chen,^{3,4} Sara Cherry,⁵ and Jeremy E. Wilusz^{1,7,*}

¹Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA

²Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences

³State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology

University of Chinese Academy of Science, Chinese Academy of Sciences, Shanghai 200031, China

⁴School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

⁵Department of Microbiology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA

⁶These authors contributed equally

⁷Lead Contact

*Correspondence: wilusz@pennmedicine.upenn.edu

<https://doi.org/10.1016/j.molcel.2017.10.034>

SUMMARY

Many eukaryotic genes generate linear mRNAs and circular RNAs, but it is largely unknown how the ratio of linear to circular RNA is controlled or modulated. Using RNAi screening in *Drosophila* cells, we identify many core spliceosome and transcription termination factors that control the RNA outputs of reporter and endogenous genes. When spliceosome components were depleted or inhibited pharmacologically, the steady-state levels of circular RNAs increased while expression of their associated linear mRNAs concomitantly decreased. Upon inhibiting RNA polymerase II termination via depletion of the cleavage/polyadenylation machinery, circular RNA levels were similarly increased. This is because readthrough transcripts now extend into downstream genes and are subjected to backsplicing. In total, these results demonstrate that inhibition or slowing of canonical pre-mRNA processing events shifts the steady-state output of protein-coding genes toward circular RNAs. This is in part because nascent RNAs become directed into alternative pathways that lead to circular RNA production.

INTRODUCTION

Upon initiation of gene transcription, a set of molecular machines is coordinately assembled on the nascent RNA, which acts to process the transcript into a mature mRNA that is exported to the cytoplasm for translation (reviewed in Bentley, 2014). On most nascent mRNAs, the capping machinery modifies the 5' end, the pre-mRNA splicing machinery removes introns, and the cleavage/polyadenylation machinery generates the mature

3' end as well as promotes transcription termination. If these steps do not occur appropriately, surveillance mechanisms are often triggered and the RNA is degraded (reviewed in Jensen et al., 2013; Kilchert et al., 2016). Surprisingly, large-scale RNA profiling efforts have indicated that >75% of the human genome is transcribed to yield a highly complex network of overlapping RNAs (Djebali et al., 2012; Hangauer et al., 2013; Wilusz et al., 2009), which includes many noncoding RNAs as well as readthrough transcripts that extend beyond annotated gene boundaries (Akiva et al., 2006; Maher et al., 2009; Parra et al., 2006; Vilborg et al., 2015). These readthrough transcripts can contain genetic information from multiple genes, although the ultimate fate of most readthrough transcripts remains unclear. One hypothesis is that some may be used for the production of stable RNA species.

Circular RNAs are a recently described class of stable RNAs that are naturally resistant to degradation by exonucleases and are generated from thousands of protein-coding genes (reviewed in Chen, 2016; Ebbesen et al., 2016; Wilusz, 2016). The mature transcripts contain almost exclusively exonic sequences, have covalently linked ends, and are generated when the pre-mRNA splicing machinery “backsplices” to join a downstream splice donor to an upstream splice acceptor (e.g., the end of exon 2 is joined to the beginning of exon 2). Backsplicing is often initiated when complementary sequences from different introns base pair and bring the intervening splice sites close together (Ivanov et al., 2015; Kramer et al., 2015; Liang and Wilusz, 2014; Zhang et al., 2014), although there are also examples in which circular RNA production is promoted by RNA-binding proteins (Ashwal-Fluss et al., 2014; Conn et al., 2015) or tied to exon-skipping events (Barrett et al., 2015; Kelly et al., 2015). While some circular RNAs are expressed at higher levels than their associated linear mRNAs, most are expressed at much lower levels (<10% of the associated linear mRNA) and accumulate to only a few molecules per cell (Guo et al., 2014; Jeck et al., 2013; Salzman et al., 2012). This suggests that backsplicing is generally inefficient in most cells (Zhang et al., 2016), and

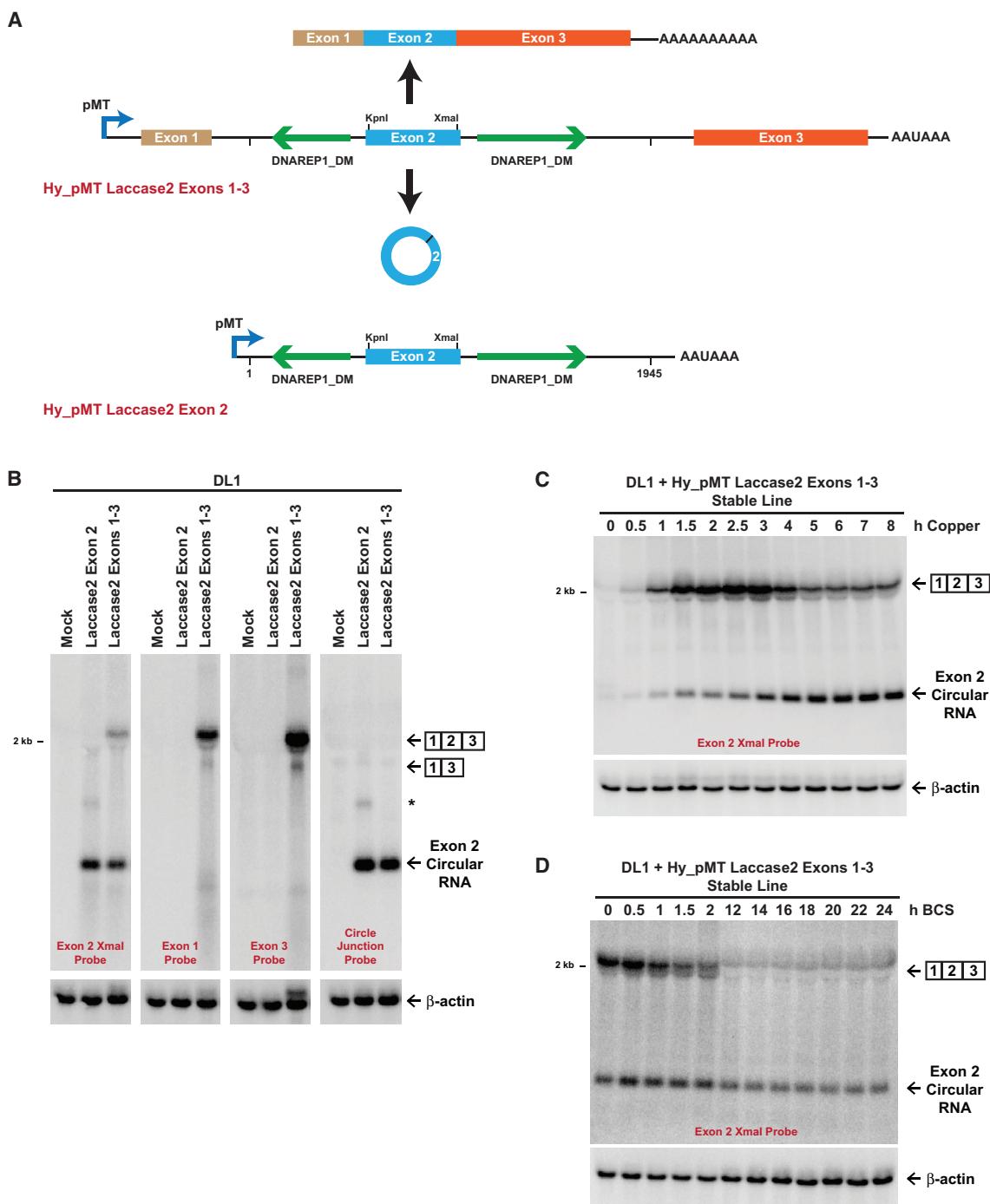


Figure 1. The Laccase2 Circular RNA Accumulates to High Levels Due to Its Long Half-Life

(A) Schematics of the Hy_pMT Laccase2 Exons 1–3 and Hy_pMT Laccase2 Exon 2 plasmids. Upon induction of the metallothionein promoter (pMT) by addition of CuSO₄, the ~3.6 kb pre-mRNA derived from the Hy_pMT Laccase2 Exons 1–3 plasmid can be canonically spliced to generate an ~1.8 kb linear mRNA that is subsequently polyadenylated (top) or backspliced to generate a 514 nt circular RNA from exon 2 (bottom). Inverted DNAREP1_DM repeat sequences (green) are present in the introns flanking exon 2. See also Figure S1A.

(B) Laccase2 plasmids were transfected into *Drosophila* DL1 cells, CuSO₄ was added for 14 hr, and northern blots were performed. *Concatenated and/or intertwined circular RNA. β-actin was used as a loading control. See also Figure S1B.

(legend continued on next page)

it remains unclear what purpose (if any) most backsplicing events serve.

At the *Drosophila Muscleblind (Mbl)* locus, backsplicing serves an auto-regulatory function and prevents Mbl protein from accumulating to excess amounts (Ashwal-Fluss et al., 2014). Given that the levels of circular RNAs and linear mRNAs from most genes are not correlated (Salzman et al., 2013), the regulatory strategy used at the *Mbl* gene seems to be a unique case. Two other circular RNAs (CDR1as/ciRS-7 and Sry) function to sponge specific microRNAs (Hansen et al., 2013; Memczak et al., 2013), but this again does not appear to be a general function for circular RNAs as most are not enriched in microRNA-binding sites (Guo et al., 2014). It has instead been suggested that circular RNAs may help form large RNA-protein complexes, e.g., at neuronal synapses (Rybäk-Wolf et al., 2015; You et al., 2015); be translated (Abe et al., 2015; Chen and Sarnow, 1995; Kramer et al., 2015; Legnini et al., 2017; Pamudurti et al., 2017; Yang et al., 2017); or be involved in innate immune responses (Chen et al., 2017; Li et al., 2017).

Given this lack of consensus about how and why circular RNAs are produced, we set out to better understand how cells modulate the relative amounts of linear versus circular RNA that are expressed from protein-coding genes. Using minigene expression plasmids based on the *Drosophila laccase2* gene, we found that canonically spliced linear mRNAs are more efficiently generated than circular RNAs. However, the circular RNAs accumulate due to their long half-lives and can ultimately become the dominant isoform observed in cells. Using RNAi screening, we surprisingly found that depletion of many core spliceosomal components, including the SF3b and SF3a complexes, resulted in specific increases in the steady-state levels of circular RNAs from reporters and endogenous *Drosophila* genes. Upon treating cells with the SF3b inhibitor Pladienolide B (Plad B), canonical linear mRNA splicing was inhibited while circular RNAs continued to accumulate. This suggests that backsplicing becomes a preferred outcome when spliceosome activity is limiting. Our RNAi screening additionally revealed that circular RNAs can be generated from readthrough transcripts that extend past annotated poly(A) signals. Upon depleting cleavage/polyadenylation factors or the torpedo exonuclease Rat1, failures in transcription termination were observed and the backsplicing machinery processed these transcripts once they extended into the downstream gene. Circular RNA expression can thus be uncoupled from its host gene promoter, a result we confirmed at the endogenous human *PAIP2* locus. In total, we propose that slowing or inhibiting canonical pre-mRNA processing events, including splicing or 3' end processing, shifts the outputs of protein-coding genes toward circular RNAs. As circular RNAs are naturally resistant to degradation by exonucleases, slight changes in the efficiency of backsplicing can result in large changes in the steady-state levels of these transcripts, thereby allowing circular RNAs to be the dominant isoform expressed from some protein-coding genes.

RESULTS

The Laccase2 Circular RNA Is Inefficiently Generated, but Accumulates Due to Its Long Half-Life

To characterize the *cis*-acting regulatory elements and *trans*-acting factors that dictate the amounts of linear and circular RNA that accumulate from protein-coding genes, we generated a three-exon minigene for the *Drosophila laccase2* gene (Figure 1A, top). This was done by modifying our previously described Hy_pMT Laccase2 Exon 2 plasmid (Figure 1A, bottom), which expresses *laccase2* exon 2 and its immediate intronic sequences under the control of the metallothionein A promoter (pMT) (Kramer et al., 2015). Exon 1 (with 187 nt of the downstream intron) and exon 3 (with 214 nt of the upstream intron) of the *laccase2* gene were added, thereby placing backsplicing in direct competition with canonical splicing (Figures 1A and S1A). Importantly, KpnI and XmaI sites were also inserted into exon 2 so that plasmid-derived transcripts could be distinguished on northern blots from endogenous Laccase2 transcripts.

The Hy_pMT Laccase2 Exon 2 and Hy_pMT Laccase2 Exons 1–3 plasmids were then transfected into *Drosophila* DL1 cells, copper (CuSO_4) was added for 14 hr to induce pMT transcription, and total RNA was isolated. Consistent with prior results (Kramer et al., 2015), northern blots revealed that the pre-mRNA derived from the Hy_pMT Laccase2 Exon 2 plasmid was efficiently backspliced to generate a 514 nt circular RNA (Figure 1B). The Hy_pMT Laccase2 Exons 1–3 plasmid likewise generated the circular RNA, which was resistant to digestion by the 3'-5' exonuclease RNase R (Figure S1B), as well as a three-exon linear mRNA (Figure 1B). An additional linear mRNA with exon 2 skipped could also be detected, but this transcript accumulated to only ~5% of the level of the fully spliced mRNA (Figure 1B). Most notably, the plasmid-derived circular RNA accumulated to higher levels than the fully spliced mRNA (Figure 1B), similar to what is observed from the endogenous *laccase2* gene in DL1 cells (Kramer et al., 2015). This three-exon minigene thus accurately recapitulates the endogenous splicing pattern, suggesting it could reveal mechanisms that regulate how much linear versus circular RNA accumulates from the *laccase2* gene.

Base pairing between highly complementary transposable elements in the flanking introns facilitates backsplicing from many eukaryotic genes (Ivanov et al., 2015; Jeck et al., 2013; Liang and Wilusz, 2014; Zhang et al., 2014). Indeed, a pair of inverted *DNAREP1_DM* family transposons flanks exon 2 of the *Drosophila laccase2* gene (Figure 1A). Prior mutational analysis showed that ~100 nt of each repeat (which are highly complementary over a ~55 nt region) is sufficient for exon circularization from the Hy_pMT Laccase2 Exon 2 plasmid (Kramer et al., 2015). To determine if the presence of flanking exons alters the extent of intronic base pairing required for backsplicing, we progressively deleted the *DNAREP1_DM* repeats from the Hy_pMT Laccase2

(C) A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1–3 plasmid was treated with CuSO_4 for the indicated amounts of time, and northern blots were performed. See also Figure S1F.

(D) To measure RNA half-lives, the stable cell line was treated with CuSO_4 for 3 hr followed by chelation of the metal with BCS for the indicated amounts of time. Cross-hybridization of the northern probe with ribosomal RNAs is responsible for the ~2 kb transcripts weakly detected after 12 hr. See also Figure S1G.

Exons 1–3 plasmid (Figure S1C). Interestingly, even when there is competition between canonical splicing and backsplicing, ~100 nt of each repeat was sufficient for the circular RNA to accumulate to high levels (Figures S1D and S1E). This supports a model in which base pairing between short complementary regions is a major driver of exon circularization.

The strong accumulation of the Laccase2 circular RNA could be due to the splicing machinery preferring to backsplice or due to differences in half-lives between the mature linear and circular RNAs. To distinguish between these models, we first took advantage of the hygromycin resistance (HygroR) cassette (Figure S1A) to generate a DL1 cell line stably expressing the inducible Hy_pMT Laccase2 Exons 1–3 plasmid. CuSO₄ was then added to induce transcription for short periods of time such that the levels of the linear and circular transcripts should largely reflect their synthesis rates (Figure 1C). Unlike what was observed at 14 hr of CuSO₄ induction (Figure 1B), the linear RNA was the predominant transcript that accumulated at early time points (e.g., 2 hr) (Figures 1C and S1F). This suggests that backsplicing is slow and not the preferred splicing outcome, a result that is consistent with recent genome-wide nascent RNA sequencing data in human cells (Zhang et al., 2016). To measure the linear and circular RNA half-lives, a 3 hr transcription pulse was induced and shut off by chelating CuSO₄ with bathocuproine disulphonate (BCS) (Djuranic et al., 2012). Whereas the linear RNA had a half-life of ~2 hr, the circular RNA remained largely stable after 24 hr (Figures 1D and S1G). These results thus indicate that the reporter circular RNA is inefficiently generated, but its long half-life allows the transcript to accumulate to levels greater than the linear mRNA.

Depleting Components of the U2 snRNP Shifts Gene Outputs toward Circular RNAs

To identify *trans*-acting factors that control the ratio of linear to circular RNA expressed from Hy_pMT Laccase2 Exons 1–3, we took advantage of the stable cell line and used double-stranded RNAs (dsRNAs) to individually knock down proteins with well-established roles in pre-mRNA splicing for 3 days. CuSO₄ was added for the last 14 hr and expression of the linear and circular RNAs was measured using northern blots (Figures 2A, 2B, and S2A). It was previously demonstrated that the SR protein SF2 (homolog of human SRSF1) inhibits backsplicing from the endogenous *laccase2* gene (Kramer et al., 2015) and thus we first asked if the Hy_pMT Laccase2 Exons 1–3 reporter recapitulated this regulation. Compared to cells treated with a control (β -galactosidase [β -gal]) dsRNA, depletion of SF2 indeed resulted in a >50-fold increase in circular RNA levels and a corresponding >4-fold decrease in spliced linear RNA levels (Figure 2A, lane 4). With these proof-of-concept experiments in hand, we then assayed the effect of depleting other pre-mRNA splicing factors, including core spliceosomal components.

Upon depleting SF3b1 (CG2807) or SF3a1 (CG16941), components of the U2 snRNP that are required for pre-spliceosome assembly (Wahl et al., 2009), we found that expression of the mature linear mRNA was largely eliminated (Figure 2A, lanes 6 and 8). Surprisingly, however, expression of the circular RNA was only slightly decreased upon SF3b1 depletion and, in fact, increased upon SF3a1 depletion (Figures 2A and 2C). The

steady-state output of the reporter gene thus became predominantly circular RNA when early steps in spliceosome assembly are slowed or inhibited. SF3b1 is the largest subunit of the seven-member SF3b protein complex (Wahl et al., 2009), and individually depleting the other members of this complex (Figure 2B, lanes 2–8) caused an increase in circular RNA levels coupled to a reduction in linear mRNA levels (Figure 2C). SF3a1 is likewise the largest subunit of the heterotrimeric SF3a complex (Lin and Xu, 2012), and depletion of SF3a2 or SF3a3 also resulted in increased levels of circular RNA (Figure 2B, lanes 10 and 11).

Upon examining expression of the endogenous *laccase2* locus by qPCR (Figures 2D) or northern blots (Figure 2E), we found that depleting the SF3b or SF3a complexes also affected endogenous splicing patterns. When any of these ten factors were depleted from DL1 cells, endogenous Laccase2 circular RNA levels increased and linear Laccase2 mRNA levels decreased (Figures 2D and 2E). To next determine if the expression of circular RNAs may generally increase when SF3b or SF3a is depleted from *Drosophila* cells, we examined the outputs of additional endogenous genes that are known to generate circular RNAs (Ashwal-Fluss et al., 2014). *Unextended* (Uex) generates a 334 nt circular RNA from exon 4 (Figure 3A) and *Plexin A* (PlexA) generates a 1,439 nt circular RNA from exon 2 (Figure 3B). Like the *laccase2* locus, complementary transposable elements are present in the introns flanking each of these exons. Depletion of SF3b or SF3a components caused significant increases in the expression of both the Uex and PlexA circular RNAs (Figures 3A and 3B).

To rule out potential off-target effects that may be caused by 3 days of dsRNA treatment, we first measured the toxicity of the various dsRNAs. Depleting SF3b6, Phf5a, or SF3a2 did not markedly affect cell viability, while depleting other SF3b and SF3a components inhibited growth to some extent (Figure S2B). Nevertheless, when we treated cells with unrelated dsRNAs (targeting unrelated genes) that caused similar levels of toxicity (Figure S2B), the relative ratio of linear versus circular RNA that accumulated from the Hy_pMT Laccase2 Exons 1–3 reporter was largely unchanged (Figure S2C). This strongly indicates that the changes in circular RNA levels that are observed upon depletion of the SF3b and SF3a complexes are not due to general toxicity of the RNAi, but because these factors are key regulators of protein-coding gene outputs.

Canonical Splicing Is More Sensitive to Pharmacological Inhibition of SF3b Than Backsplicing

To further determine if the changes in the steady-state amounts of linear versus circular RNA are due to differences in their biogenesis efficiencies, we treated the Hy_pMT Laccase2 Exons 1–3 stable cell line for 6 hr with CuSO₄ and Plad B, a naturally occurring anti-tumor drug that binds and inhibits SF3b (Kotake et al., 2007). When cells were treated with 5 or 10 nM Plad B, linear reporter mRNAs failed to accumulate, while circular RNAs were strikingly still produced (Figure 3C). As the dose of Plad B was further increased, circular RNA production decreased, indicating that some amount of active SF3b is required for backsplicing. Nevertheless, circular RNA production is clearly more resistant to SF3b inhibition than canonical

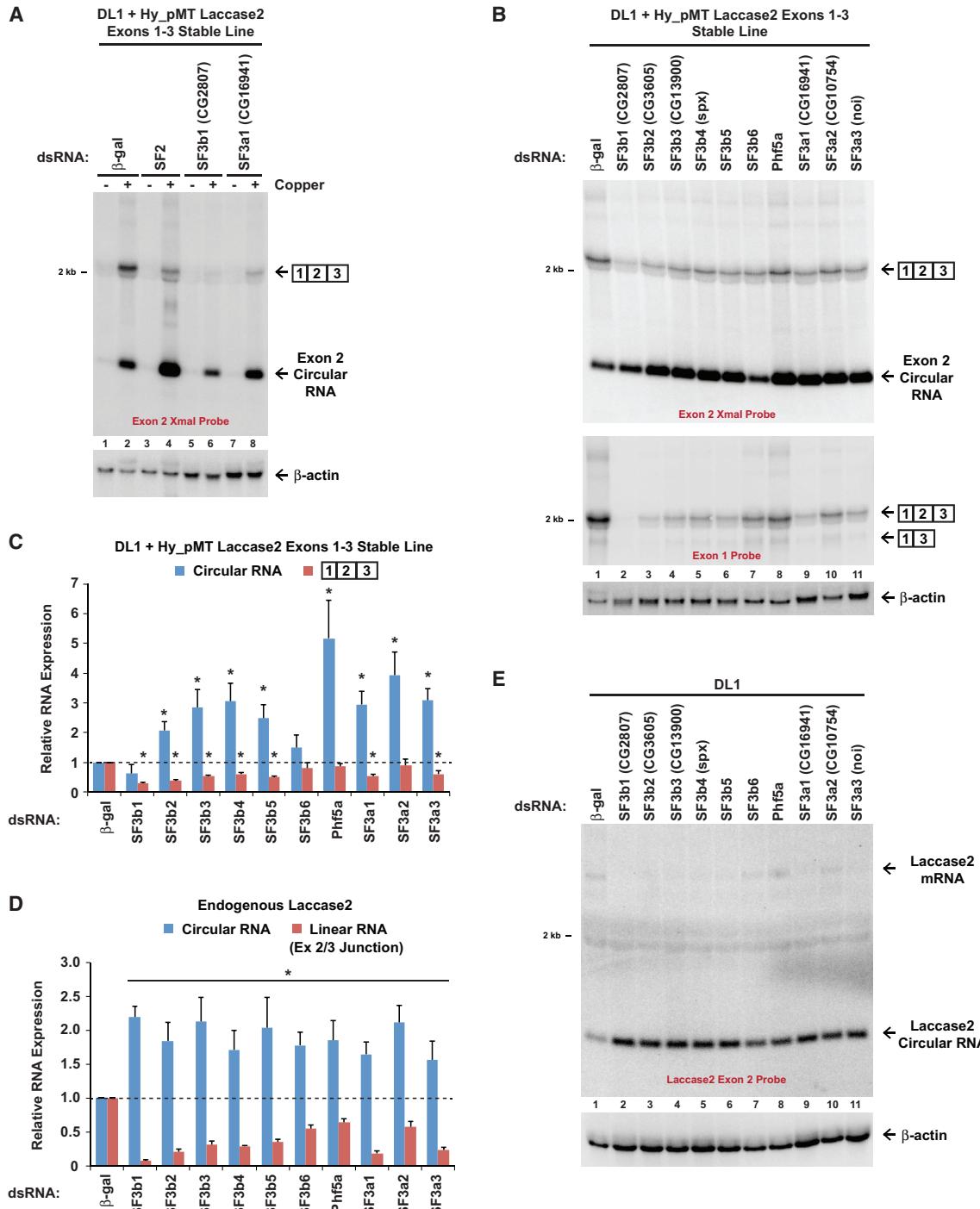


Figure 2. The SF3b and SF3a Complexes Regulate Laccase2 Circular RNA Levels

(A and B) A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1–3 plasmid was treated with the indicated dsRNAs for 3 days and, where noted, CuSO₄ was added for the last 14 hr. Northern blots were used to examine expression of reporter-derived transcripts. Representative blots are shown. See also Figure S2.

(C) Linear mRNA and circular RNA levels were quantified using ImageQuant from three independent northern blot experiments. Data are normalized to the β-gal dsRNA samples and are shown as mean ± SD. *p < 0.01.

(D) DL1 cells were treated with the indicated dsRNAs for 3 days. qRT-PCR was then used to quantify expression of the endogenous Laccase2 linear mRNA (exon 2/3 junction) and Laccase2 circular RNA. Data from four independent experiments were normalized to β-actin (Act42a) and are shown as mean ± SD. *p < 0.01.

(E) Northern blots were used to measure expression of the endogenous laccase2 gene in DL1 cells that had been treated with dsRNAs for 3 days.

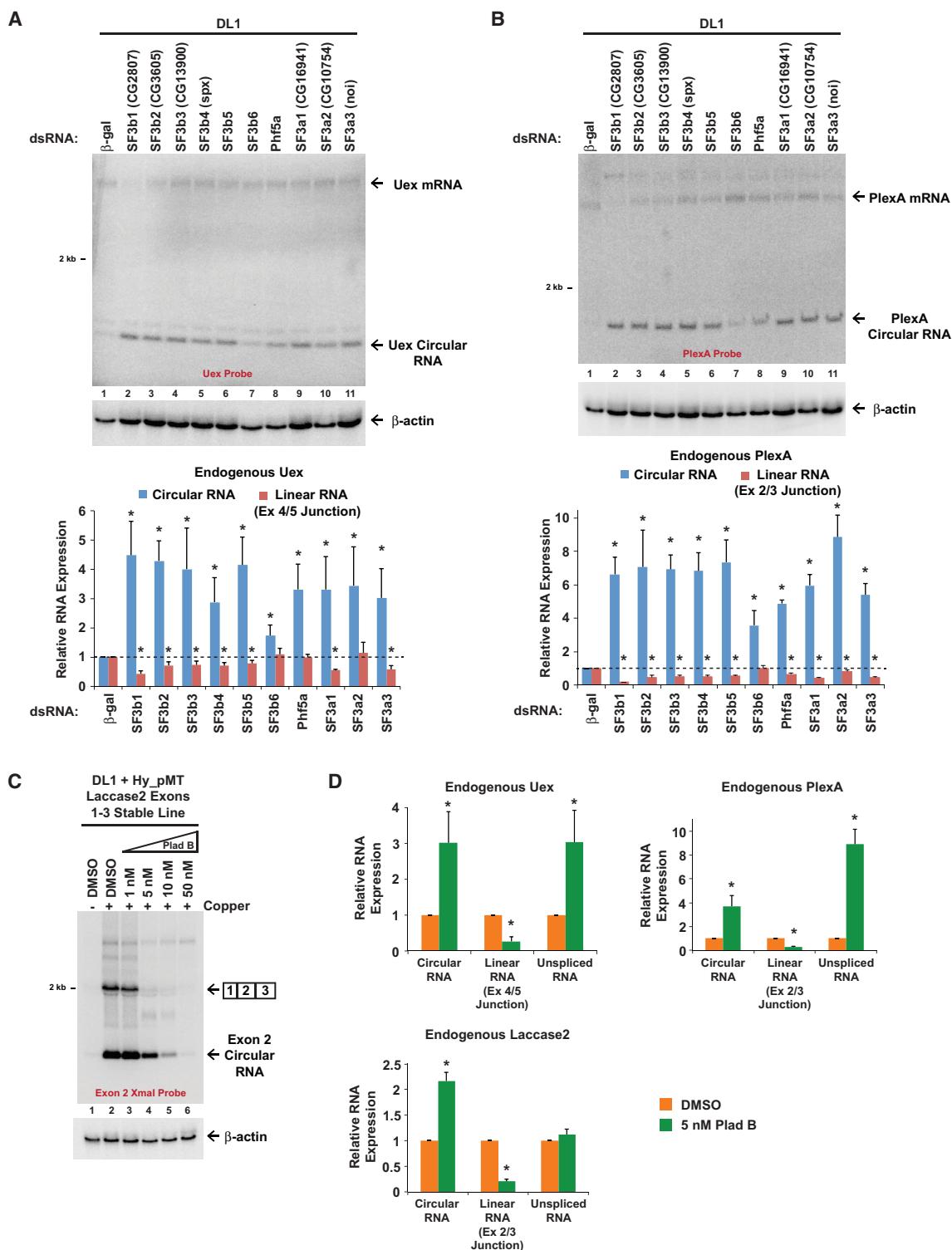


Figure 3. Depleting or Pharmacologically Inhibiting SF3b Results in Increased Expression of Endogenous Circular RNAs

(A) DL1 cells were treated with the indicated dsRNAs for 3 days. Northern blots (top) and qRT-PCR (bottom) were then used to quantify expression of the endogenous Uex linear mRNA and circular RNA. A representative blot is shown. Data from four independent qRT-PCR experiments were normalized to β -actin (Act42a) and are shown as mean \pm SD. *p < 0.05.

(legend continued on next page)

splicing. To determine if endogenous backsplicing events are similarly resistant to Plad B treatment, we treated DL1 cells with 5 nM Plad B for 6 hr and measured the outputs of the endogenous *PlexA*, *Uex*, and *laccase2* genes (Figure 3D). Compared to control-treated cells, the levels of each of the linear mRNAs were all significantly reduced by Plad B treatment (Figure 3D). In contrast, expression of each of the endogenous circular RNAs increased by 2- to 4-fold. These data thus mirror the RNAi results and suggest that the relative efficiency of backsplicing increases when spliceosome assembly is slowed or when spliceosomes are in limited quantities.

Endogenous Circular RNA Levels Increase upon Depleting a Variety of Core Spliceosomal Components

As the spliceosome is formed by a dynamic assembly of five small nuclear RNAs (snRNAs) and ~170 proteins (reviewed in Wahl et al., 2009), we next addressed if depletion of other core spliceosomal components would similarly affect circular RNA levels. For this analysis, we focused on the endogenous *PlexA* gene because the *PlexA* circular RNA was most affected by depletion of SF3b and SF3a components (Figure 3B). Thirty-five proteins that function at different stages of spliceosome assembly or catalysis (Mount and Salz, 2000) were individually depleted from DL1 cells, and northern blots (Figure 4A) or qPCR (Figure S3) was used to measure the output of the *PlexA* gene. Remarkably, depletion of 25 of these core spliceosomal proteins resulted in >2-fold increases in circular RNA expression. A number of factors with very different annotated functions caused greater than 10-fold increases, including (1) snRNP-U1-70K and snRNP-U1-C, which are involved in 5' splice site recognition; (2) Prp8, the largest protein component in the spliceosome, which crosslinks with U5 and U6 snRNAs as well as all the critical sites of chemistry (the 5' splice site, branch point sequence, and 3' splice site); and (3) Slu7 and CG6015 (CDC40), which act in the second catalytic step of splicing. This thus indicates that changing the levels of many core spliceosomal components can have a profound impact on circular RNA levels.

For the core spliceosomal components that caused the largest increases in *PlexA* circular RNA levels, we then asked if these factors similarly regulate other endogenous circular RNAs (Figure 4B). In general, depleting core spliceosomal components resulted in global increases in circular RNA levels, although the effect sizes varied depending on the factor that was depleted and the circular RNA examined. For example, depletion of snRNP-U1-70K or snRNP-U1-C caused *PlexA* and *pasilla* (ps) circular RNA levels to increase >10-fold, but had a smaller effect on expression of the *Laccase2*, *Uex*, *minibrain* (mnb), CG42663, and ectoderm-expressed 4 (Ect4) circular RNAs (Figure 4B).

Consistent with these data, modulating core spliceosomal components has been previously demonstrated to affect only a subset of canonical splicing events in *Drosophila* (Brooks et al., 2015; Park et al., 2004). In total, our data suggest a model in which backsplicing becomes a preferred mode of pre-mRNA processing when spliceosomal components are in limited quantities, thereby causing the steady-state output of genes to be biased toward circular RNAs.

Circular RNAs Can Be Generated from Readthrough Transcription Events

Beyond regulation of circular RNA levels by the splicing machinery, recent work suggests that backsplicing may often occur post-transcriptionally after RNA polymerase II has reached the end of the parent gene (Liang and Wilusz, 2014; Zhang et al., 2016). This led us to predict that inhibiting co-transcriptional 3' end processing may increase circular RNA levels. Indeed, depleting the pre-mRNA 3' end processing endonuclease Cpsf73 (Figure 5A, lane 4; Figure S4A) caused expression of the reporter circular RNA to increase almost 50-fold while linear RNA levels decreased by ~2-fold (Figure 5B). We initially suspected that Cpsf73 was acting at the 3' end of the *laccase2* minigene, and thus we replaced the downstream SV40 poly(A) signal with a self-cleaving hammerhead ribozyme (HhRz) sequence (Figure S4B). Surprisingly, depletion of Cpsf73 caused Laccase2 circular RNA levels to increase even when the HhRz was present (Figure S4C).

Further work showed that depletion of Cpsf73 caused the reporter-derived circular RNA to be expressed at a high level even when CuSO₄ was not added (Figure 5A, lane 3; Figure 5B), suggesting that the circular RNA was transcribed from a constitutive promoter. As depletion of Cpsf73 did not cause the endogenous pMT to be active in the absence of CuSO₄ (Figure S4D), we reasoned that the reporter-derived circular RNA was likely not being transcribed from its expected promoter. To rule out the existence of a cryptic promoter within the *laccase2* minigene, this sequence was replaced with a portion of the *Drosophila* datti pre-mRNA that can generate a circular RNA independently of inverted intronic repeats (Figure S4E). Expression of the datti circular RNA could be induced upon CuSO₄ treatment in control-treated cells, but Cpsf73 depletion caused the reporter-derived datti circular RNA to be constitutively produced (Figure S4E). We thus concluded that a constitutive promoter upstream of pMT was likely driving expression of the Laccase2 and datti circular RNAs from the reporters.

Located upstream is a constitutive copia transposon LTR promoter that drives expression of the HygroR gene (Figures 5C and S1A). In control-treated DL1 cells that stably maintain the Hy_pMT Laccase2 Exons 1–3 plasmid, the hygromycin

(B) DL1 cells were treated as in (A). Northern blots (top) and qRT-PCR (bottom) were used to quantify expression of the endogenous *PlexA* linear mRNA and circular RNA. A representative blot is shown. Data from four independent qRT-PCR experiments were normalized to β-actin (Act42a) and are shown as mean ± SD. *p < 0.05.

(C) A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1–3 plasmid was treated with increasing concentrations of Pladienolide B (Plad B) and, where noted, CuSO₄ for 6 hr. Northern blots were used to examine expression of reporter-derived transcripts.

(D) DL1 cells were treated with DMSO or 5 nM Plad B for 6 hr. Total RNA was then isolated and subjected to qRT-PCR to measure the expression of transcripts derived from the *PlexA*, *Uex*, and *laccase2* genes. Data from three independent experiments were normalized to β-actin (Act42a) and are shown as mean ± SD. *p < 0.05.

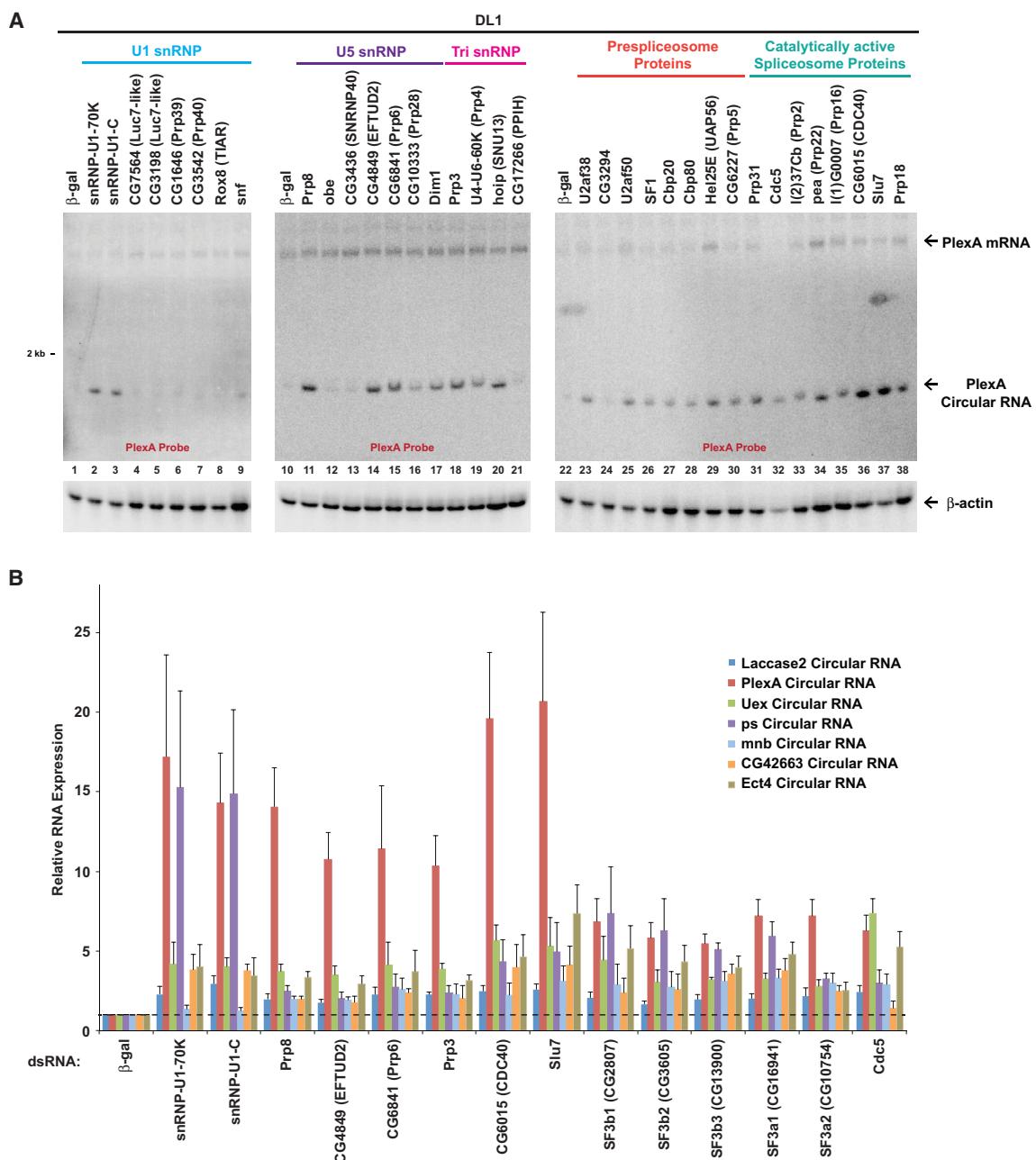


Figure 4. Depleting Many Core Spliceosomal Components Results in Increased Expression of Endogenous Circular RNAs

(A) DL1 cells were treated with the indicated dsRNAs for 3 days and northern blots were used to quantify expression of the endogenous PlexA linear mRNA and circular RNA. Core spliceosomal components that function at different stages of spliceosome assembly were individually depleted. See also Figure S3.

(B) Total RNA was isolated from DL1 cells treated with dsRNAs for 3 days and subjected to qRT-PCR to measure the expression of endogenous circular RNAs. Data from four independent experiments were normalized to β -actin (β -Act42a) and are shown as mean \pm SD.

mRNA is constitutively transcribed and efficiently cleaved/polyadenylated using the SV40 poly(A) signal at its 3' end (Figure 5C, top). As *laccase2* minigene transcripts do not appreciably accumulate when CuSO_4 is absent (Figures 2A and 5A), RNA polymerase II derived from the copia promoter rarely (if ever) transcribes far into the downstream gene. However, when Cpsf73 is depleted, we reasoned that the HygroR pre-mRNA

fails to be cleaved at its 3' end, resulting in transcriptional readthrough into the downstream *laccase2* minigene (Figure 5C, bottom). Once both inverted repeats have been transcribed, the intervening exon can be backspliced to generate a circular RNA. As predicted by this model, depletion of Symplekin (Symp), a scaffolding protein for the cleavage/polyadenylation machinery, also resulted in generation of the circular RNA (Figure 5D).

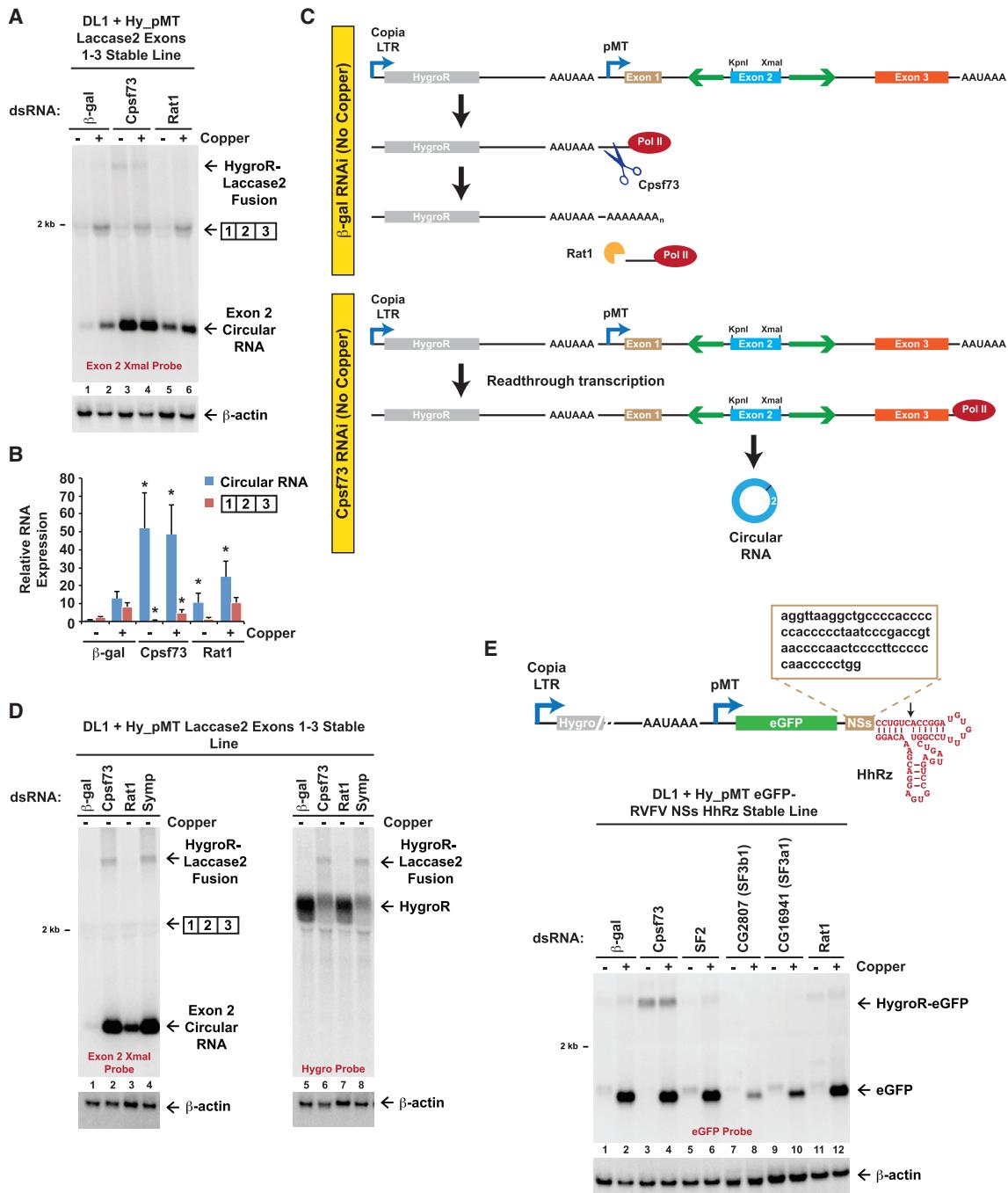


Figure 5. Readthrough Transcription from the Upstream HygroR Gene Enables Circular RNA Production from the Minigene Reporter

(A) A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1–3 plasmid was treated with the indicated dsRNAs for 3 days and, where noted, CuSO₄ was added for the last 14 hr. Northern blots were used to examine expression of reporter-derived transcripts. See also Figure S4A.

(B) Linear mRNA and circular RNA levels were quantified using ImageQuant from four independent northern blot experiments and were normalized to the “β-gal, No copper” samples. Data are shown as mean ± SD. *p < 0.05.

(C) Schematic of readthrough transcription model. In control-treated cells (top), only the upstream copia transposon LTR promoter is active when no CuSO₄ is present. This results in the production of the HygroR mRNA, which is cleaved/polyadenylated, and the downstream RNA is rapidly degraded by the Rat1 exonuclease to facilitate transcription termination. When Cpsf73 is depleted (bottom), 3' processing of the HygroR mRNA is inefficient, resulting in readthrough transcription into the downstream *laccase2* minigene. When the intronic repeats (green) base pair to one another, a circular RNA is produced. See also Figure S4.

(D) DL1 cells stably expressing the Hy_pMT Laccase2 Exons 1–3 plasmid were treated with the indicated dsRNAs for 3 days and total RNA was isolated. Northern blots using probes complementary to Laccase2 exon 2 (lanes 1–4) or the HygroR mRNA (lanes 5–8) were performed.

(legend continued on next page)

Furthermore, we detected a >4 kb linear transcript that was produced regardless of whether CuSO₄ was added when Cpsf73 or Symp was depleted (Figure 5A). This transcript corresponds to a fusion HygroR-Laccase2 linear mRNA that has been polyadenylated downstream of the *laccase2* minigene (Figure 5D).

The torpedo model of transcription termination posits that cleavage by Cpsf73 provides an entry site for a 5'-3' exonuclease (Rat1 in *Drosophila*), which degrades the downstream RNA in order to catch up to the elongation complex and cause it to terminate (Connelly and Manley, 1988). Consistent with the torpedo model, we found that depletion of Rat1 caused expression of the reporter circular RNA to become constitutive (Figures 5A, lanes 5 and 6; Figure 5B). A fusion HygroR-Laccase2 linear mRNA did not accumulate upon Rat1 depletion, indicating that Rat1 functions only after 3' cleavage (Figures 5A and 5D).

As a final test of the readthrough transcription model, we replaced the *laccase2* minigene with an eGFP open reading frame followed by the self-cleaving HhRz (Figure 5E, top). Because ribozyme cleavage is not sufficient to stabilize the 3' end of a mature linear mRNA (Figure S4C), we added a 74 nt sequence from the 3' end of the Rift Valley fever virus (RVFV) NSs mRNA immediately upstream of the HhRz sequence (Moleston et al., 2016) (Figure 5E, top). This NSs sequence is able to block 3'-5' degradation, allowing the non-polyadenylated eGFP mRNA to accumulate to high levels in cells (Figure 5E, lane 2). After generating a stable DL1 cell line expressing the eGFP plasmid, we predicted that Cpsf73 depletion would result in generation of a fusion HygroR-eGFP linear mRNA. Indeed, an ~3,500 nt transcript was detected in cells depleted of Cpsf73, irrespective of CuSO₄ addition (Figure 5E, lanes 3 and 4). These data demonstrate that depletion of Cpsf73, Symp, or Rat1 causes readthrough transcription from the HygroR gene, which can result in the production of stable circular RNAs and fusion linear RNAs. This suggests that circular RNAs do not have to be made from their host gene promoters and that modulating transcription termination efficiencies can alter the steady-state levels of circular RNAs.

Production of Endogenous Circular RNAs by Readthrough Transcription

Having established that transcription termination defects can lead to circular RNA accumulation from reporters, we asked if endogenous circular RNAs are produced from readthrough transcription events. Human embryonic carcinoma PA1 cells were treated with 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) for 3 hr to arrest transcription, followed by pulse labeling with the uridine analog 4-thiouridine (4sU) and purification of newly synthesized RNAs at different time points after DRB removal (Zhang et al., 2016) (Figure 6A). Thirty-nine potential polycistronic transcripts (indicative of readthrough transcription) were detected upon quantifying expression in intergenic regions between gene pairs (Figures S5A and S5B). We focused our validation efforts on the human *MATR3* and *PAIP2* genes, which are

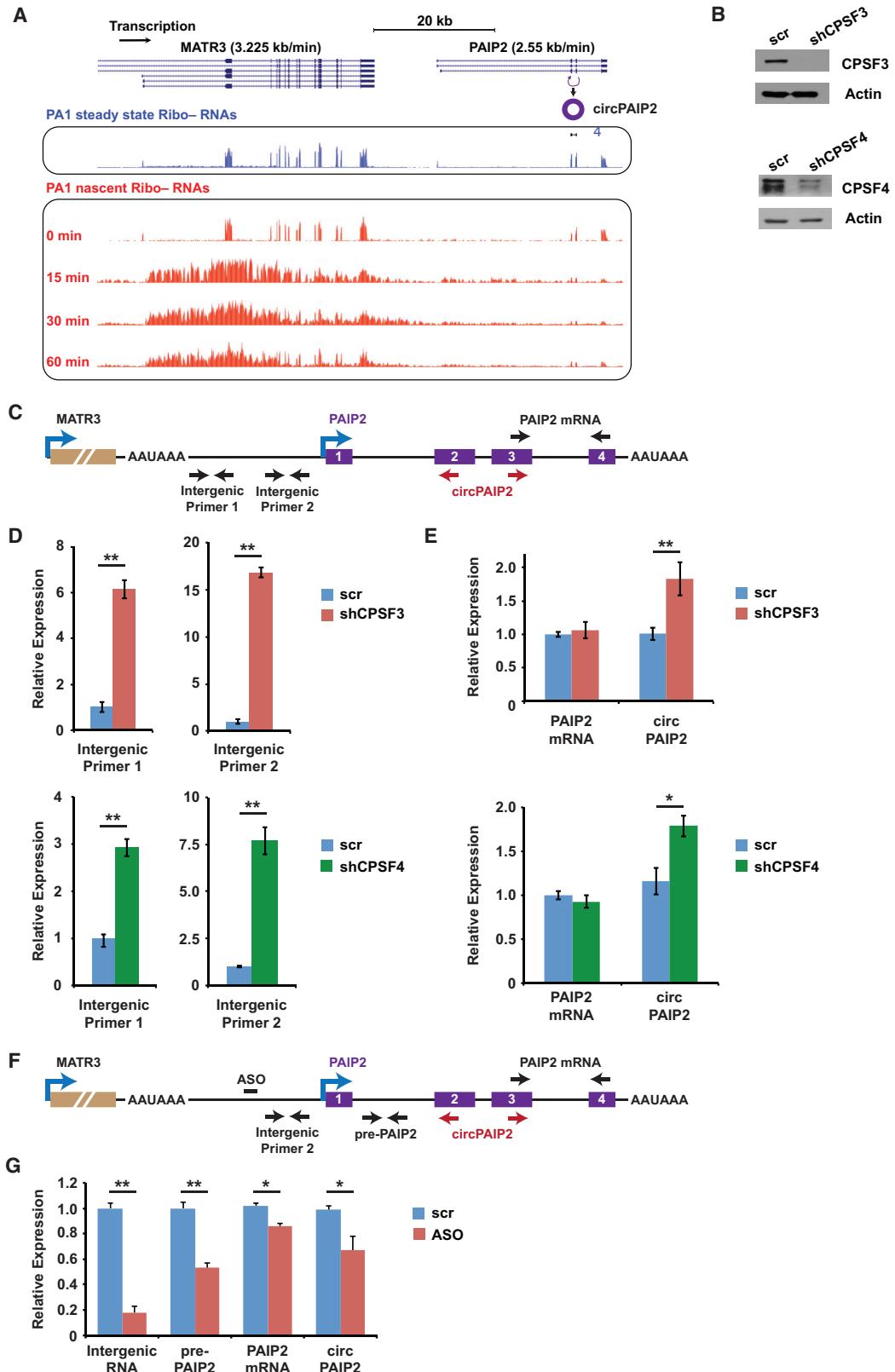
in the same transcriptional orientation, with the downstream *PAIP2* gene generating a circular RNA (circPAIP2) from exons 2 and 3 (Figures 6A and S5C). The transcription elongation rate of the *MATR3* gene is faster than average in PA1 cells (Zhang et al., 2016), and significant amounts of nascent RNAs were continuously detected throughout the ~10 kb intergenic region between *MATR3* and *PAIP2* (Figures 6A and S5C). This suggested that *MATR3* primary transcripts may often extend into the downstream *PAIP2* gene, and some of these RNAs may be backspliced to yield the *PAIP2* circular RNA.

Short hairpin RNAs (shRNAs) were first used to deplete CPSF3 (CPSF73) or CPSF4 (CPSF30) from PA1 cells (Figure 6B) and qRT-PCR used to quantify RNAs derived from the *MATR3-PAIP2* intergenic region or the *PAIP2* gene (Figure 6C). Consistent with a defect in *MATR3* transcription termination, depletion of CPSF3 or CPSF4 resulted in increased expression of transcripts from the *MATR3-PAIP2* intergenic region (Figure 6D). Expression of circPAIP2 also increased, suggesting that these *MATR3* readthrough transcripts are likely used to generate stable circular RNAs from the downstream *PAIP2* gene (Figure 6E). To directly confirm that readthrough transcription from *MATR3* is required for the increase in circular RNA levels, PA1 cells were treated with a phosphorothioate-modified antisense oligonucleotide (ASO) that targeted the *MATR3-PAIP2* intergenic region (Figure 6F). Compared to cells treated with a scrambled (scr) oligo, ASO treatment reduced intergenic RNA levels by ~80% and simultaneously caused significant decreases in the levels of the *PAIP2* pre-mRNA, linear mRNA, and circular RNA (Figure 6G). Of note, ASO treatment reduced *PAIP2* circular RNA levels by >30%, indicating that production of this endogenous circular RNA is often not from its host gene promoter and instead from a readthrough transcript. In total, these results indicate that failures to co-transcriptionally cleave at the 3' ends of genes, such as human *MATR3*, can result in the production of circular RNAs from downstream genes.

DISCUSSION

Thousands of circular RNAs can be generated from protein-coding regions, but we are only beginning to understand how, when, and why these RNAs are made. In the current study, we demonstrated that backsplicing is normally less efficient than canonical splicing, but the steady-state levels of *Drosophila* circular RNAs significantly increase upon inhibition or slowing of multiple canonical pre-mRNA processing events. While linear mRNA levels were decreased upon depletion of many core spliceosomal components, circular RNAs became remarkably enriched under these conditions (Figure 7A). Some of this effect is due to the long half-lives of circular RNAs, but our data also indicate that backsplicing becomes a preferred splicing outcome when active spliceosomes are limiting (Figures 2, 3, and 4). Inhibition of 3' end processing and transcription termination likewise resulted in increased circular RNA levels (Figures 5 and 6). This is because

(E) The *laccase2* minigene from the Hy_pMT Laccase2 Exons 1–3 plasmid was replaced with an eGFP open reading frame. In addition, the downstream polyadenylation signal was replaced with the hammerhead ribozyme (HhRz) preceded by a 74 nt sequence from the 3' end of the Rift Valley fever virus (RVFV) NSs mRNA (brown). A DL1 cell line stably expressing this plasmid was then generated and treated with the indicated dsRNAs for 3 days. Depletion of Cpsf73 resulted in the production of a fusion HygroR-eGFP mRNA that is stabilized at its 3' end by the RVFV NSs sequence.



(legend on next page)

circular RNA biogenesis can occur on readthrough transcripts that extend into downstream genes (Figure 7B). Altogether, these results suggest that inhibition or slowing of co-transcriptional processing events allows increased opportunity for back-splicing reactions to occur. As mature circular RNAs have long half-lives, they are able to accumulate to high levels and become the predominant output of some genes.

Canonical Splicing and Backsplicing Have Different Sensitivities to Spliceosome Inhibition

As a nascent RNA is being transcribed, 5' and 3' splice sites at the ends of exons are recognized and spliceosomes begin to be assembled in a stepwise manner. Depending on what splice sites are paired together, a variety of mature RNAs can be produced, including circular RNAs. In many cases, pre-mRNA splicing outcomes are dictated by *cis*-acting sequence elements (namely, intronic and exonic splicing enhancers or silencers) that serve as binding sites for *trans*-acting factors, such as SR proteins or hnRNPs, that act positively or negatively to recruit snRNPs to nearby splice sites (reviewed in Busch and Hertel, 2012; Smith and Valcárcel, 2000). Indeed, depletion of SR proteins or hnRNPs results in significant changes to alternative splicing patterns (Brooks et al., 2015), including effects on circular RNA levels (Figure 2A) (Kramer et al., 2015). In the current work, we surprisingly demonstrated that circular RNA levels also significantly increase upon alterations in the amounts or activities of core spliceosomal components (Figures 2, 3, and 4). Modulating the levels of spliceosomal components has previously been shown to affect individual splicing decisions in yeast (Pleiss et al., 2007), *Drosophila* (Brooks et al., 2015; Park et al., 2004), and mammalian cells (Papasaikas et al., 2015; Saltzman et al., 2011), but the effects were often quite small and gene specific. In contrast, we demonstrated here that the expression of circular RNAs from multiple endogenous loci increased when spliceosomal components were depleted (Figure 4B). The long half-lives of circular RNAs likely magnify the changes in biogenesis efficiencies, resulting in the large effect sizes observed.

Considering that both canonical splicing and backsplicing require splice sites (Starke et al., 2015), why do these processes display such distinct sensitivities to spliceosome inhibition? We propose that this difference may be due, at least in part, to the process by which exonic sequences are first identified in most pre-mRNAs (Berget, 1995). During exon definition, initial splicing complexes are formed across exons: U1 and U2 snRNPs bind at opposite ends of each exon, stabilized by additional factors, such as SR proteins, that establish a network of protein-protein interactions (Figure 7A, left). These cross-exon inter-

actions must then somehow be converted to cross-intron interactions (thereby pairing splice sites from separate exons) in order to yield a canonically spliced linear mRNA. However, to generate a circular RNA, these initial exon definition complexes likely do not have to be disrupted and, in fact, may be stabilized by base pairing between complementary sequences in the flanking introns (Figure 7A, right). Furthermore, these exon definition complexes may help directly promote backsplicing as they can contain the U4/U6.U5 tri-snRNP and potentially be converted to B-like pre-catalytic splicing complexes (Schneider et al., 2010). We thus propose that when core spliceosomal components are depleted from cells, cross-exon interactions may be slowly or rarely converted into productive cross-intron interactions. This allows catalytically competent backsplicing complexes to be assembled more frequently, thus resulting in production of more circular RNAs and fewer canonical linear mRNAs (Figure 7A).

It remains possible that backsplicing only requires a subset of the factors that are normally required for canonical linear splicing (or, conversely, that backsplicing requires factors that are dispensable for canonical linear splicing), but further detailed mechanistic studies are required to clarify this point. Regardless of the exact molecular details, the current work clearly demonstrates that backsplicing and canonical splicing are differentially affected when core spliceosomal components are inhibited or depleted from cells. Our study thus predicts that endogenous tissues expressing limiting amounts of spliceosomal components should produce higher levels of circular RNAs. It will be particularly informative to determine if this model holds true *in vivo* in neuronal or aging tissues, which produce the highest amounts of annotated circular RNAs (Rybäk-Wolf et al., 2015; Westholm et al., 2014; You et al., 2015).

Readthrough Transcription Can Be Exploited to Generate Circular RNAs from Downstream Regions

RNA polymerase II sometimes precisely terminates after reaching a poly(A) signal, but termination at most genes occurs at variable distances (sometimes as much as 10 kb) downstream of the poly(A) signal (Proudfoot, 2016; Shi and Manley, 2015). This allows many RNAs to be generated from intergenic regions (Vilborg et al., 2015) and, in fact, the global efficiency of termination is known to change in cancer (Grosso et al., 2015; Kannan et al., 2011; Maher et al., 2009), viral infection (Nemeroff et al., 1998; Rutkowski et al., 2015), and osmotic stress (Vilborg et al., 2015). Once produced, some readthrough transcripts have been proposed to help maintain the nuclear scaffold during osmotic stress (Vilborg et al., 2015). Others serve as precursors

Figure 6. Readthrough Transcription Generates Endogenous Circular RNAs from the Human PAIP2 Gene

- (A) Readthrough transcription from MATR3 as shown by nascent RNA sequencing in human PA1 cells (Zhang et al., 2016). The downstream PAIP2 gene generates a circular RNA from exons 2 and 3 (purple). See also Figure S5.
- (B) Immunoblot to confirm depletion of CPSF3 (top) or CPSF4 (bottom) by shRNAs in PA1 cells. scr, scrambled shRNA control. Actin was used as a loading control.
- (C) Schematic showing qRT-PCR primers used to quantify transcripts from the MATR3-PAIP2 locus.
- (D and E) qRT-PCR quantification of transcripts from the MATR3-PAIP2 intergenic region (D) or the PAIP2 gene (E) in PA1 cells depleted of CPSF3 or CPSF4. Data are shown as mean \pm SD from three independent experiments. * p < 0.05, ** p < 0.01.
- (F) PA1 cells were treated for 8 hr with a phosphorothioate-modified antisense oligonucleotide (ASO) complementary to the MATR3-PAIP2 intergenic region.
- (G) qRT-PCR was then used to quantify transcripts from the MATR3-PAIP2 intergenic region or the PAIP2 gene. Locations of primers are shown in (F). Data are shown as mean \pm SD from three independent experiments. * p < 0.05, ** p < 0.01.

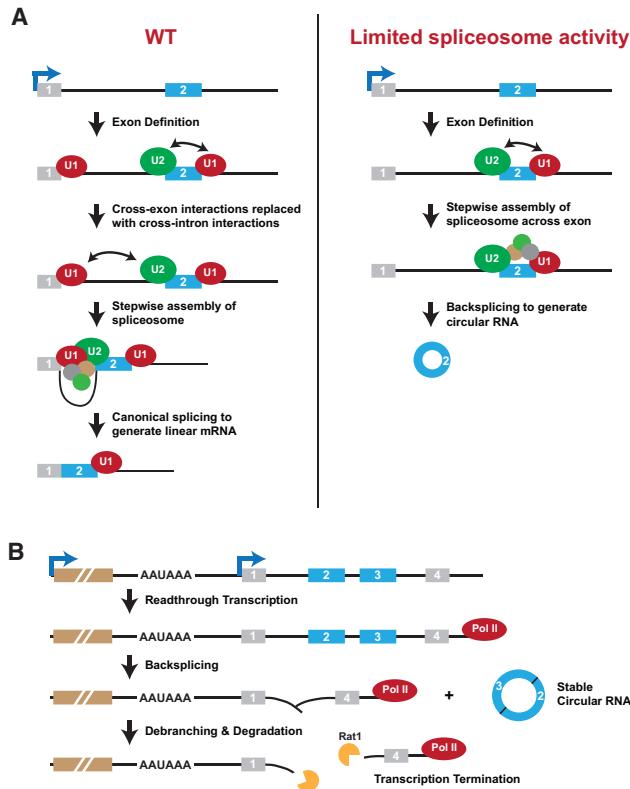


Figure 7. Proposed Models for How Inhibition or Slowing of Canonical Pre-mRNA Processing Events Can Result in Increased Circular RNA Levels

(A) In wild-type (WT) cells (left), exons within pre-mRNAs are first defined and spliceosomal components assemble across each exon. U1 snRNP recognizes the downstream 5' splice site, U2 snRNP binds the upstream polypyrimidine tract and branch point sequence, and factors such as SR proteins mediate cross-exon interactions. These cross-exon interactions are subsequently replaced with cross-intron interactions to enable full assembly of the spliceosome and generation of a linear mRNA. When spliceosome activity is limiting (e.g., due to depletion of core spliceosome components) (right), we propose that cross-exon interactions are not easily replaced with cross-intron interactions. The full spliceosome thus assembles across an exon, resulting in backsplicing and the generation of a circular RNA.

(B) Failure to efficiently terminate transcriptional units (e.g., due to depletion of cleavage/polyadenylation factors) can cause nascent RNAs to be extended into downstream genes. If the appropriate signals are present in this readthrough transcript (such as inverted intronic repeats), backsplicing can occur to release a mature circular RNA. The remaining nascent RNA likely consists of a Y-shaped structure with a 2'-5' phosphodiester bond at the upstream branch site. This structure may be debranched, thereby providing an entry site for exonucleases, including Rat1, that enable transcription termination.

to stable RNAs, including *Drosophila* Piwi-interacting RNAs (piRNAs) (Chen et al., 2016; Mohn et al., 2014) and human SPA long noncoding RNAs (Wu et al., 2016) that lack their own promoter sequences.

In the current study, we found that circular RNAs can be efficiently generated from readthrough transcription in a manner highly analogous to how piRNAs and SPA RNAs are generated. Once both flanking intronic repeats have been transcribed, backsplicing can occur, thereby physically separating the

mature circular RNA from the nascent transcript (Figure 7B). Such backsplicing events may be fairly common in wild-type cells as we found that at least 30% of the endogenous PAIP2 circular RNA is made from readthrough transcription from the upstream *MATR3* gene (Figure 6). On the other hand, readthrough transcription can interfere with the activation of downstream promoters, and it is thus likely that the expression of a number of circular RNAs will decrease when the global efficiency of termination is altered.

Once generated, mature circular RNAs likely have their own unique functions (Hansen et al., 2013; Memczak et al., 2013), but we further suggest that the backsplicing events responsible for their biogenesis may also serve a role in aiding in transcription termination. This is because backsplicing likely leaves a Y-shaped structure with a 2'-5' phosphodiester bond at the branch site in the remaining nascent RNA (Figure 7B). The lariat debranching enzyme may act upon this structure, thereby creating an entry site for Rat1 or other 5'-3' exonucleases to terminate downstream transcription. It will be very informative in the future to determine if backsplicing events truly help repress gene transcription levels as well as prevent polymerases from continuing to transcribe unchecked.

In summary, our findings provide key insights into how the spliceosome and pre-mRNA 3' end processing machineries regulate gene outputs and dictate the relative amounts of linear versus circular RNA that accumulate from protein-coding genes. In particular, we demonstrated that inhibition or slowing of pre-mRNA processing events can profoundly increase the steady-state levels of circular RNAs, in part by increasing their biogenesis. We thus propose that the global differences in circular RNA levels that are observed across cell types (Westholm et al., 2014) may be due to differences in the concentrations or activities of pre-mRNA processing components. When such factors are limiting, canonical pre-mRNA processing events may not occur quickly, thereby allowing nascent RNAs to be directed toward alternative pathways that lead to circular RNA production. As factors like SF3b1 and U2AF1 (homolog of *Drosophila* U2af38) are commonly mutated in multiple human cancers (Yoshida et al., 2011), an attractive hypothesis is that these mutant proteins may be altering the amounts of linear versus circular RNAs made from protein-coding genes, thereby resulting in disease phenotypes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Cell lines
 - Microbe strains
- METHOD DETAILS
 - Expression plasmid construction
 - *Drosophila* cell culture and transfections
 - *Drosophila* RNAi
 - Mammalian cell culture and transfections

- Northern blotting
- qRT-PCR
- Computational identification of polycistronic transcripts from nascent RNA-seq

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, and supplemental methods and can be found with this article online at <https://doi.org/10.1016/j.molcel.2017.10.034>.

AUTHOR CONTRIBUTIONS

J.E.W. conceived and designed the project. D.L., D.C.T., and J.E.W. designed and performed experiments with help from S.C.; Z.L. and L.Y. performed nascent RNA sequencing bioinformatics analyses; H.W. and L.-L.C. designed and performed *PAIP2* experiments. D.L., D.C.T., and J.E.W. wrote the paper with input from the authors.

ACKNOWLEDGMENTS

We thank Gideon Dreyfuss and all members of the Wilusz, Chen, Cherry, and Yang labs for discussions and advice. This work was supported by NIH grants R00-GM104166, R35-GM119735, and R01-NS099371; start-up funds from the University of Pennsylvania; and grants 91540115, 91440202, and 31471241 from NSFC. J.E.W. is a Rita Allen Foundation Scholar. S.C. is a recipient of the Burroughs Wellcome Investigators in the Pathogenesis of Infectious Disease Award.

Received: March 23, 2017

Revised: July 31, 2017

Accepted: October 26, 2017

Published: November 22, 2017

REFERENCES

- Abe, N., Matsumoto, K., Nishihara, M., Nakano, Y., Shibata, A., Maruyama, H., Shuto, S., Matsuda, A., Yoshida, M., Ito, Y., and Abe, H. (2015). Rolling circle translation of circular RNA in living human cells. *Sci. Rep.* 5, 16435.
- Akiva, P., Toporik, A., Edelheit, S., Peretz, Y., Diber, A., Shemesh, R., Novik, A., and Sorek, R. (2006). Transcription-mediated gene fusion in the human genome. *Genome Res.* 16, 30–36.
- Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evental, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014). circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* 56, 55–66.
- Barrett, S.P., Wang, P.L., and Salzman, J. (2015). Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *eLife* 4, e07540.
- Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* 15, 163–175.
- Berget, S.M. (1995). Exon recognition in vertebrate splicing. *J. Biol. Chem.* 270, 2411–2414.
- Brooks, A.N., Duff, M.O., May, G., Yang, L., Bolisetty, M., Landolin, J., Wan, K., Sandler, J., Booth, B.W., Celtniker, S.E., et al. (2015). Regulation of alternative splicing in *Drosophila* by 56 RNA binding proteins. *Genome Res.* 25, 1771–1780.
- Busch, A., and Hertel, K.J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA* 3, 1–12.
- Chen, L.L. (2016). The biogenesis and emerging roles of circular RNAs. *Nat. Rev. Mol. Cell Biol.* 17, 205–211.
- Chen, C.Y., and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* 268, 415–417.
- Chen, Y.A., Stuwe, E., Luo, Y., Ninova, M., Le Thomas, A., Rozhavskaya, E., Li, S., Vempati, S., Laver, J.D., Patel, D.J., et al. (2016). Cutoff suppresses RNA polymerase II termination to ensure expression of piRNA precursors. *Mol. Cell* 63, 97–109.
- Chen, Y.G., Kim, M.V., Chen, X., Batista, P.J., Aoyama, S., Wilusz, J.E., Iwasaki, A., and Chang, H.Y. (2017). Sensing self and foreign circular RNAs by intron identity. *Mol. Cell* 67, 228–238.e5.
- Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A., and Goodall, G.J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell* 160, 1125–1134.
- Connelly, S., and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* 2, 440–452.
- Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108.
- Djurjanovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* 336, 237–240.
- Ebbesen, K.K., Kjems, J., and Hansen, T.B. (2016). Circular RNAs: identification, biogenesis and function. *Biochim. Biophys. Acta* 1859, 163–168.
- Grosso, A.R., Leite, A.P., Carvalho, S., Matos, M.R., Martins, F.B., Vitor, A.C., Desterro, J.M., Carmo-Fonseca, M., and de Almeida, S.F. (2015). Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. *eLife* 4, e09214.
- Guo, J.U., Agarwal, V., Guo, H., and Bartel, D.P. (2014). Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* 15, 409.
- Hangauer, M.J., Vaughn, I.W., and McManus, M.T. (2013). Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet.* 9, e1003569.
- Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388.
- Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H.T., Orejuela, M.R., Piechotta, M., Levanon, E.Y., Landthaler, M., Dieterich, C., and Rajewsky, N. (2015). Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* 10, 170–177.
- Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F., and Sharpless, N.E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19, 141–157.
- Jensen, T.H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. *Mol. Cell* 52, 473–484.
- Kannan, K., Wang, L., Wang, J., Ittmann, M.M., Li, W., and Yen, L. (2011). Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proc. Natl. Acad. Sci. USA* 108, 9172–9177.
- Kelly, S., Greenman, C., Cook, P.R., and Papantonis, A. (2015). Exon skipping is correlated with exon circularization. *J. Mol. Biol.* 427, 2414–2417.
- Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* 17, 227–239.
- Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat. Chem. Biol.* 3, 570–575.
- Kramer, M.C., Liang, D., Tatomer, D.C., Gold, B., March, Z.M., Cherry, S., and Wilusz, J.E. (2015). Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev.* 29, 2168–2182.
- Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., Fatica, A., Santini, T., Andronache, A., Wade, M., et al. (2017). Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis. *Mol. Cell* 66, 22–37.e9.
- Li, X., Liu, C.X., Xue, W., Zhang, Y., Jiang, S., Yin, Q.F., Wei, J., Yao, R.W., Yang, L., and Chen, L.L. (2017). Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. *Mol. Cell* 67, 214–227.e7.

- Liang, D., and Wilusz, J.E. (2014). Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.* 28, 2233–2247.
- Lin, P.C., and Xu, R.M. (2012). Structure and assembly of the SF3a splicing factor complex of U2 snRNP. *EMBO J.* 31, 1579–1590.
- Maher, C.A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., and Chinnaiyan, A.M. (2009). Transcriptome sequencing to detect gene fusions in cancer. *Nature* 458, 97–101.
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersten, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333–338.
- Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* 157, 1364–1379.
- Molleston, J.M., Sabin, L.R., Moy, R.H., Menghani, S.V., Rausch, K., Gordesky-Gold, B., Hopkins, K.C., Zhou, R., Jensen, T.H., Wilusz, J.E., and Cherry, S. (2016). A conserved virus-induced cytoplasmic TRAMP-like complex recruits the exosome to target viral RNA for degradation. *Genes Dev.* 30, 1658–1670.
- Mount, S.M., and Salz, H.K. (2000). Pre-messenger RNA processing factors in the *Drosophila* genome. *J. Cell Biol.* 150, F37–F44.
- Nemeroff, M.E., Barabino, S.M., Li, Y., Keller, W., and Krug, R.M. (1998). Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Mol. Cell* 1, 991–1000.
- Pamudurti, N.R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., Hanan, M., Wyler, E., Perez-Hernandez, D., Ramberger, E., et al. (2017). Translation of circRNAs. *Mol. Cell* 66, 9–21.e7.
- Papasaikas, P., Tejedor, J.R., Vigevani, L., and Valcárcel, J. (2015). Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol. Cell* 57, 7–22.
- Park, J.W., Parisky, K., Celotto, A.M., Reenan, R.A., and Graveley, B.R. (2004). Identification of alternative splicing regulators by RNA interference in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 101, 15974–15979.
- Parra, G., Reymond, A., Dabbs, N., Dermizakis, E.T., Castelo, R., Thomson, T.M., Antonarakis, S.E., and Guigó, R. (2006). Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res.* 16, 37–44.
- Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol.* 5, e90.
- Proudfoot, N.J. (2016). Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. *Science* 352, aad9926.
- Rutkowski, A.J., Erhard, F., L'Hernault, A., Bonfert, T., Schilhabel, M., Crump, C., Rosenstiel, P., Efstratiou, S., Zimmer, R., Friedel, C.C., and Dölen, L. (2015). Widespread disruption of host transcription termination in HSV-1 infection. *Nat. Commun.* 6, 7126.
- Rybáková, A., Stottmeister, C., Glážar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* 58, 870–885.
- Saltzman, A.L., Pan, Q., and Blencowe, B.J. (2011). Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev.* 25, 373–384.
- 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.
- Salzman, J., Chen, R.E., Olsen, M.N., Wang, P.L., and Brown, P.O. (2013). Cell-type specific features of circular RNA expression. *PLoS Genet.* 9, e1003777.
- Schneider, M., Will, C.L., Anokhina, M., Tazi, J., Urlaub, H., and Lührmann, R. (2010). Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes. *Mol. Cell* 38, 223–235.
- Shi, Y., and Manley, J.L. (2015). The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev.* 29, 889–897.
- Smith, C.W., and Valcárcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.* 25, 381–388.
- Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L.H., and Bindereif, A. (2015). Exon circularization requires canonical splice signals. *Cell Rep.* 10, 103–111.
- Tatomer, D.C., Liang, D., and Wilusz, J.E. (2017). Inducible expression of eukaryotic circular RNAs from plasmids. *Methods Mol. Biol.* 1648, 143–154.
- Vilborg, A., Passarelli, M.C., Yario, T.A., Tykowski, K.T., and Steitz, J.A. (2015). Widespread inducible transcription downstream of human genes. *Mol. Cell* 59, 449–461.
- Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–718.
- Westholm, J.O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S.E., Graveley, B.R., and Lai, E.C. (2014). Genome-wide analysis of *Drosophila* circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* 9, 1966–1980.
- Wilusz, J.E. (2016). Circular RNAs: unexpected outputs of many protein-coding genes. *RNA Biol.* 1–11.
- Wilusz, J.E., Sunwoo, H., and Spector, D.L. (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* 23, 1494–1504.
- Wu, H., Yin, Q.F., Luo, Z., Yao, R.W., Zheng, C.C., Zhang, J., Xiang, J.F., Yang, L., and Chen, L.L. (2016). Unusual processing generates SPA lncRNAs that sequester multiple RNA binding proteins. *Mol. Cell* 64, 534–548.
- Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L.L., Wang, Y., et al. (2017). Extensive translation of circular RNAs driven by N(6)-methyladenosine. *Cell Res.* 27, 626–641.
- Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 478, 64–69.
- You, X., Vlatkovic, I., Babic, A., Will, T., Epstein, I., Tushev, G., Akbalik, G., Wang, M., Glock, C., Quedenau, C., et al. (2015). Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* 18, 603–610.
- Zhang, X.O., Wang, H.B., Zhang, Y., Lu, X., Chen, L.L., and Yang, L. (2014). Complementary sequence-mediated exon circularization. *Cell* 159, 134–147.
- Zhang, Y., Xue, W., Li, X., Zhang, J., Chen, S., Zhang, J.L., Yang, L., and Chen, L.L. (2016). The biogenesis of nascent circular RNAs. *Cell Rep.* 15, 611–624.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CPSF3	Bethyl	Cat#A301-091A
CPSF4	Bethyl	Cat#A301-584A
β-Actin	Sigma	Cat#A3854
Bacterial and Virus Strains		
One Shot TOP10 <i>E. coli</i> competent cells	Thermo Fisher Scientific	Cat#C404006
Chemicals, Peptides, and Recombinant Proteins		
Effectene	QIAGEN	Cat#301427
Copper sulfate	Fisher BioReagents	Cat#BP346-500
Trizol	Thermo Fisher Scientific	Cat#15596018
Bathocuproine disulphonate	Sigma	Cat#B1125-500MG
Pladienolide B	EMD Millipore Calbiochem	Cat#5.30196.0001
RNase R	Epicentre	Cat#RNR07250
Critical Commercial Assays		
MEGAscript Kit	Thermo Fisher Scientific	Cat#AMB13345
DNA-free Kit	Thermo Fisher Scientific	Cat#AM1906
Superscript III	Thermo Fisher Scientific	Cat#18080051
NorthernMax Formaldehyde Load Dye	Thermo Fisher Scientific	Cat#AM8552
NorthernMax 10X Running Buffer	Thermo Fisher Scientific	Cat#AM8671
NorthernMax 10X Denaturing Gel Buffer	Thermo Fisher Scientific	Cat#AM8676
ULTRAhyb-Oligo	Thermo Fisher Scientific	Cat#AM8663
Deposited Data		
RNA-seq of human PA1 nascent RNAs	Zhang et al., 2016	GEO: GSE73325
Experimental Models: Cell Lines		
DL1	Dr. Sara Cherry	N/A
PA1	ATCC	Cat#CRL-1572
Oligonucleotides		
Primers	This paper	Table S1
ASOs	This paper	Table S1
Recombinant DNA		
Plasmid: Hy_pMT Laccase2 Exon 2	Kramer et al., 2015	https://www.addgene.org/69890/
Plasmid: Hy_pMT Laccase2 Exons 1-3	This paper	https://www.addgene.org/91799/
Plasmid: Hy_pMT Laccase2 Exons 1-3 Δ300	This paper	N/A
Plasmid: Hy_pMT Laccase2 Exons 1-3 Δ400	This paper	N/A
Plasmid: Hy_pMT Laccase2 Exons 1-3 Δ450	This paper	N/A
Plasmid: Hy_pMT Laccase2 Exons 1-3 Δ500	This paper	N/A
Plasmid: Hy_pMT Laccase2 Exons 1-3 450-1245	This paper	N/A
Plasmid: Hy_pMT dati	This paper	https://www.addgene.org/91800/
Plasmid: Hy_pMT Laccase2 Exons 1-3 HhRz	This paper	N/A
Software and Algorithms		
CIRCExplorer	Zhang et al., 2014	https://github.com/YangLab/CIRCExplorer
Identification of polycistronic transcripts	This paper	N/A
ImageQuant	GE Healthcare	Cat#29000605

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Amaxa Nucleofector II Device	Lonza	Cat#AAD-1001S
Fisherbrand Electroporation Cuvettes 2mm Gap	Fisher Scientific	Cat#FB102

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jeremy E. Wilusz (wilusz@pennmedicine.upenn.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

Drosophila DL1 cells were grown at 25°C in Schneider's *Drosophila* medium (Thermo Fisher Scientific 21720024), supplemented with 1% (v/v) penicillin-streptomycin (Thermo Fisher Scientific 15140122), 1% (v/v) L-glutamine (Thermo Fisher Scientific 35050061), and 10% (v/v) fetal bovine serum (HyClone SH30910.03). DL1 stable cell lines were generated and maintained by selection with 150 µg/mL hygromycin B.

PA1 cells were grown at 37°C, 5% CO₂ in MEMa supplemented with 10% (v/v) fetal bovine serum, 1% (v/v) L-glutamine, and 1% (v/v) penicillin-streptomycin.

Microbe strains

One Shot TOP10 *E. coli* competent cells (Thermo Fisher Scientific C404006) were stored at -80°C and grown in LB medium at 37°C.

METHOD DETAILS

Expression plasmid construction

To generate expression plasmids for transfection into *Drosophila* cells, the indicated sequences were inserted into a pMK33/pMtHy-based plasmid between the metallothionein promoter (pMT) and the SV40 polyadenylation signal as further described in [Methods S1](#).

***Drosophila* cell culture and transfections**

For transient transfections into *Drosophila* DL1 cells, 2 × 10⁶ cells were plated in complete media in 6-well dishes and 2 µg of each expression plasmid was transfected using Effectene (QIAGEN 301427; 16 µL Enhancer and 8 µL Effectene Reagent). On the following day, a final concentration of 500 µM copper sulfate (Fisher BioReagents BP346-500) was added for 14 h (where indicated) to induce transcription from the metallothionein promoter. Total RNA was then isolated using Trizol (Thermo Fisher Scientific 15596018) as per the manufacturer's instructions.

To measure RNA half-lives in DL1 cells stably expressing the Hy_pMT Laccase2 Exons 1-3 plasmid, 500,000 cells/well were seeded in 12-well dishes and incubated for 2 days. 500 µM copper sulfate was added for 3 h followed by washing the cells twice with media containing 500 µM bathocuproine disulphonate (BCS; Sigma B1125-500MG). New media containing 50 µM BCS was then added and samples collected at the indicated time points using Trizol.

Pladienolide B (EMD Millipore Calbiochem 5.30196.0001) was added to cells for 6 h and total RNA isolated using Trizol.

***Drosophila* RNAi**

Double-stranded RNAs from the DRSC (*Drosophila* RNAi Screening Center) were generated by *in vitro* transcription (MEGAscript kit, Thermo Fisher Scientific AMB13345) of PCR templates containing the T7 promoter sequence on both ends. Primer sequences are provided in [Table S1](#). Knockdown experiments in 12-well dishes were then performed by bathing 500,000 cells with 2 µg of dsRNA. Cells were incubated for 3 days and a final concentration of 500 µM copper sulfate was added for the final 14 h (where indicated). Total RNA was isolated using Trizol.

To quantify toxicity caused by the dsRNAs, cells were re-suspended 2 h prior to collection and 10% of the sample was re-plated (50 µl/well, 3 wells per dsRNA) in 96-well plates (Corning 3904). After 2 h, cells were fixed in 5% formaldehyde and counterstained with Hoechst 33342 (Sigma B2261) to visualize nuclei. Four images per well were captured at 20x using an automated microscope (ImageXpress Micro, Molecular Devices) and analyzed with MetaXpress software to measure total cells. The average cell number for each set of images was calculated and normalized to β-gal dsRNA treated cells.

Mammalian cell culture and transfections

Metabolic labeling of nascent RNAs in human cells with 4sU and RNA-seq analysis were previously described (Zhang et al., 2016). shRNAs to deplete CPSF3 or CPSF4 from PA1 cells and western blot procedure were previously described (Wu et al., 2016). Phosphorothioate-modified antisense oligonucleotides (ASOs) were synthesized at BioSune (Shanghai, China) and introduced to PA1 cell by nucleofection (Lonza) according to the manufacturer's protocol. After 8 h ASO treatment, total RNA was isolated using Trizol and analyzed by qRT-PCR. ASO sequences are provided in Table S1.

Northern blotting

Northern blots using NorthernMax reagents (Thermo Fisher Scientific) and oligonucleotide probes were performed as previously described (Tatomer et al., 2017). All oligonucleotide probe sequences are provided in Table S1. Blots were viewed and quantified with the Typhoon 9500 scanner (GE Healthcare) and quantified using ImageQuant (GE Healthcare). Representative blots are shown. For RNase R treatments, 20 µg of total RNA was treated with 10 U RNase R (Epicentre RNR07250) at 37°C for 10 min.

qRT-PCR

Total RNAs were isolated using Trizol (Thermo Fisher Scientific 15596018), treated with DNase I (DNA-free Kit, Thermo Fisher Scientific AM1906), and cDNA was reversed transcribed using SuperScript III (Thermo Fisher Scientific 18080051). For all qRT-PCRs, actin mRNA was used as an internal control to normalize the data. qRT-PCR primer sequences are provided in Table S1.

Computational identification of polycistronic transcripts from nascent RNA-seq

Circular RNAs were identified by CIRCExplorer (Zhang et al., 2014) and potential polycistronic transcripts identified as follows. All adjacent gene pairs that have the same transcriptional direction (with no additional annotated genes in between) were first selected from RefSeq. Gene pairs without a circRNA derived from the downstream gene were then filtered out. The remaining interval regions were split into bins of 500 bp with 50 bp overhangs to generate a bin queue I.

$$I = [B_1, B_2, \dots, B_p]$$

RPKM values from the RNA-seq were used to define the expression level of each bin and the longest continuous queue Q was defined from the interval bins.

$$Q = [b_1, b_2, \dots, b_k] (b \in B)$$

Bins in the queue Q were required to meet the following two conditions:

- (1) Expression of each bin in Q is greater than 1.

$$\text{Exp}(b_i) \geq 1, i \text{ from } 1 \text{ to } k.$$

- (2) There is no significant expression changes between two adjacent bins from the queue Q.

$$0.5 < \text{Exp}(b_i)/\text{Exp}(b_{i-1}) < 2, i \text{ from } 2 \text{ to } k.$$

The continuity of the interval region was then estimated by calculating the CI (Continuity Index).

$$CI = k/p$$

Gene pairs with CI greater than 0.5 were considered as potential polycistronic transcripts.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical significance for comparisons of means was assessed by Student's t test for qRT-PCRs and Northern blots. Statistical details and error bars are defined in each figure legend.

Supplemental Information

The Output of Protein-Coding Genes

Shifts to Circular RNAs When the Pre-mRNA

Processing Machinery Is Limiting

Dongming Liang, Deirdre C. Tatomer, Zheng Luo, Huang Wu, Li Yang, Ling-Ling Chen, Sara Cherry, and Jeremy E. Wilusz

Figure S1

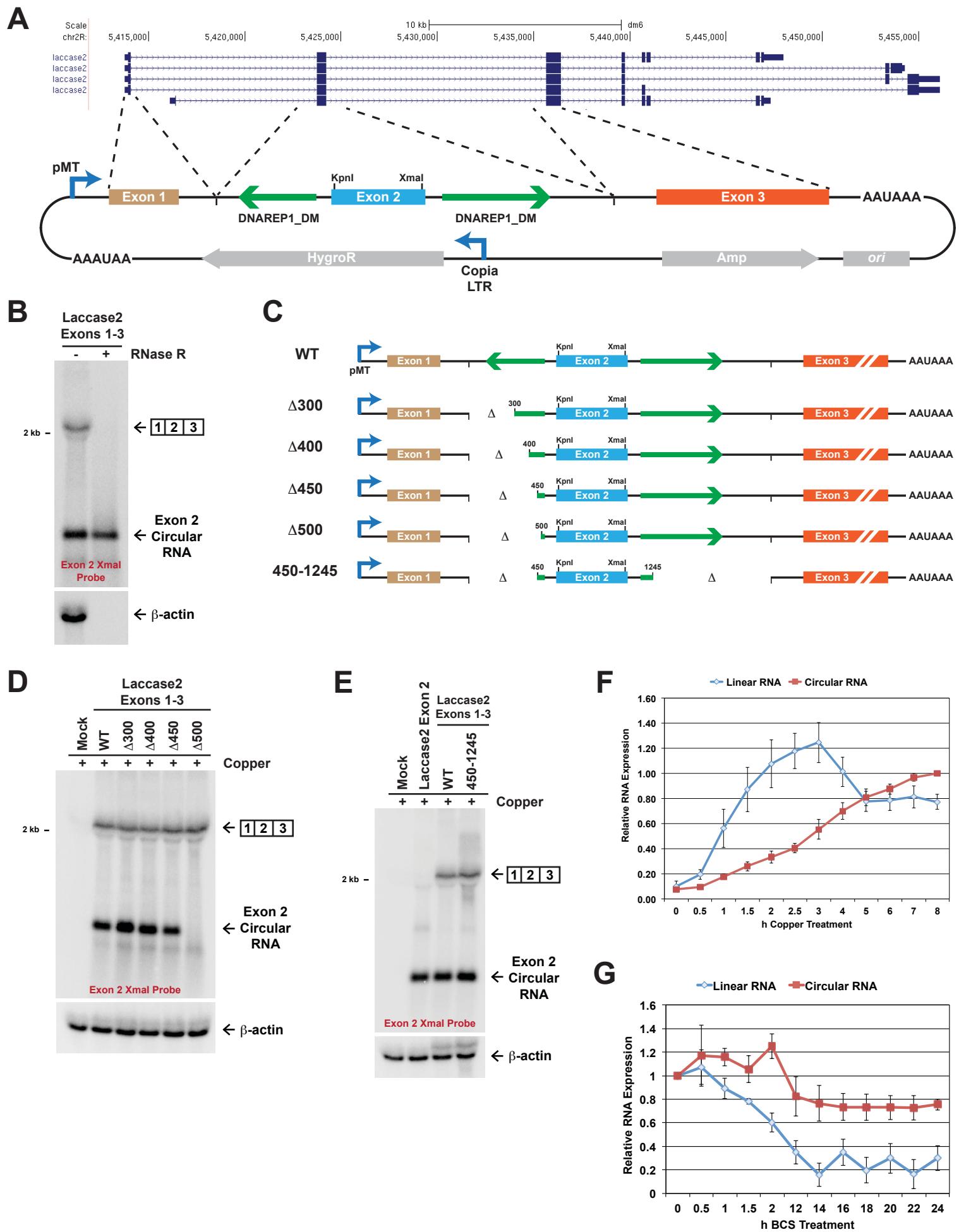


Figure S2

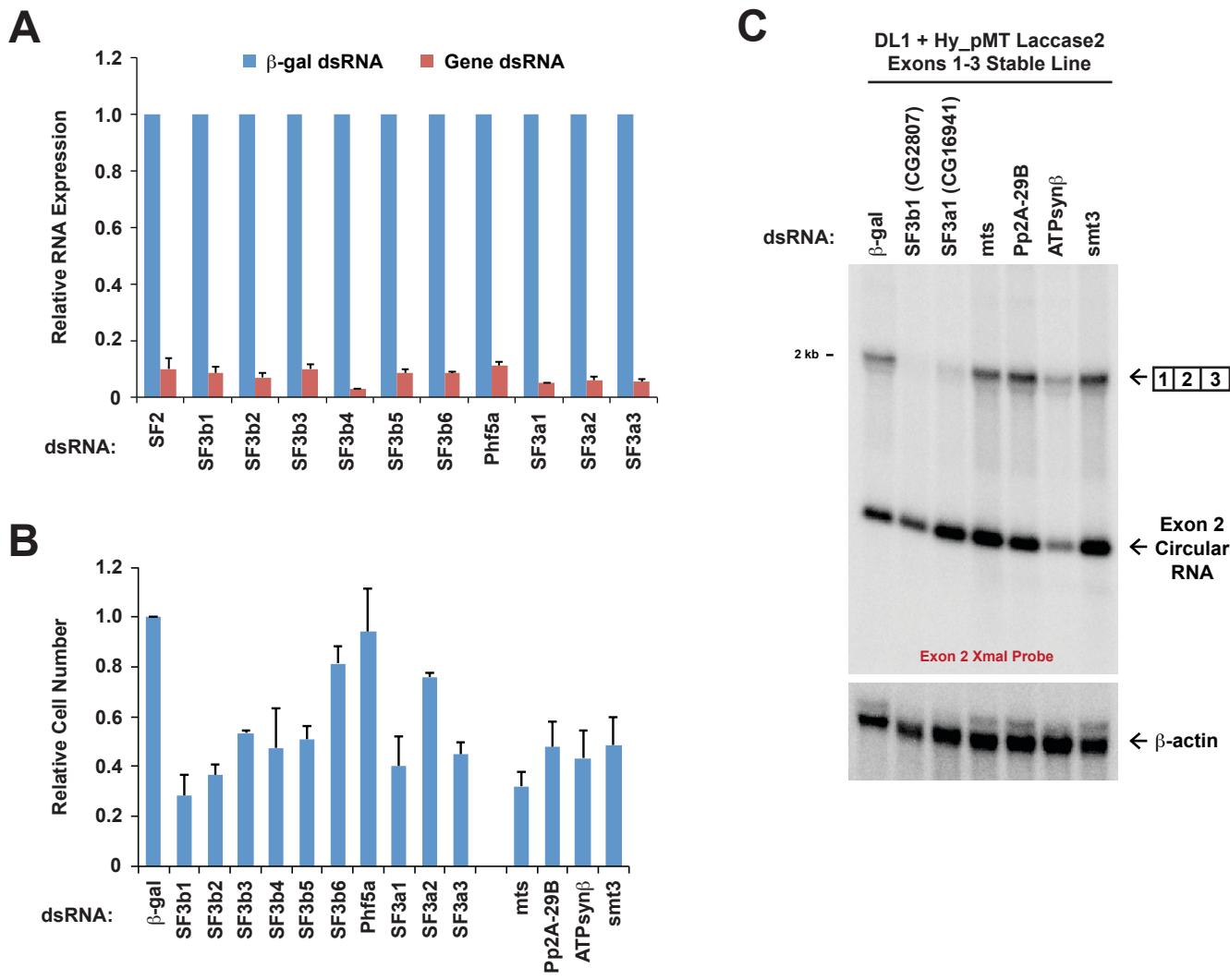


Figure S3

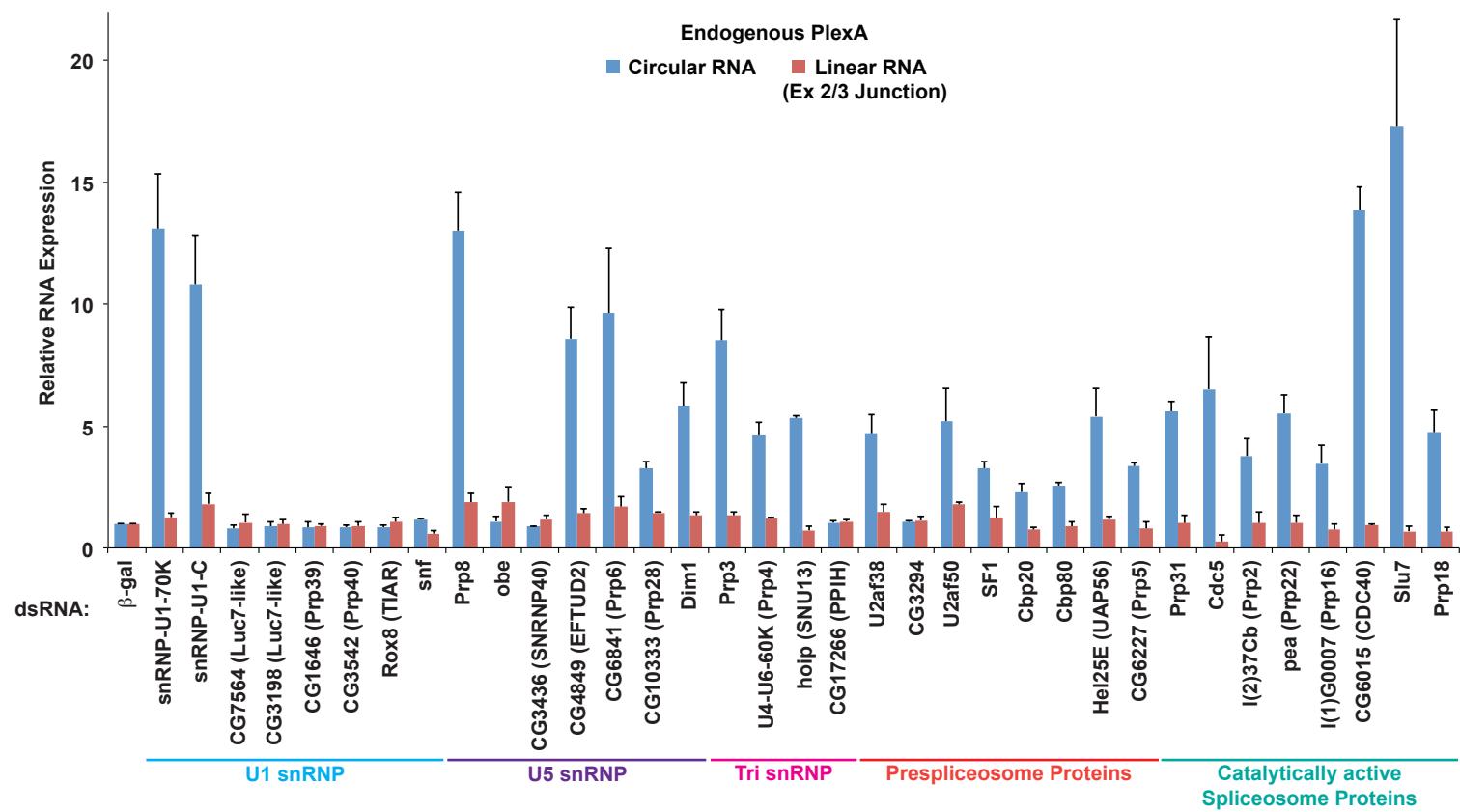


Figure S4

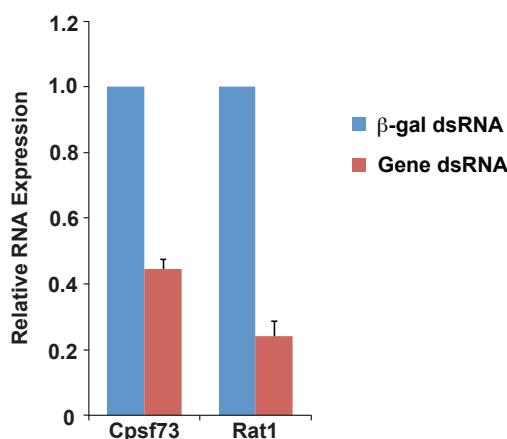
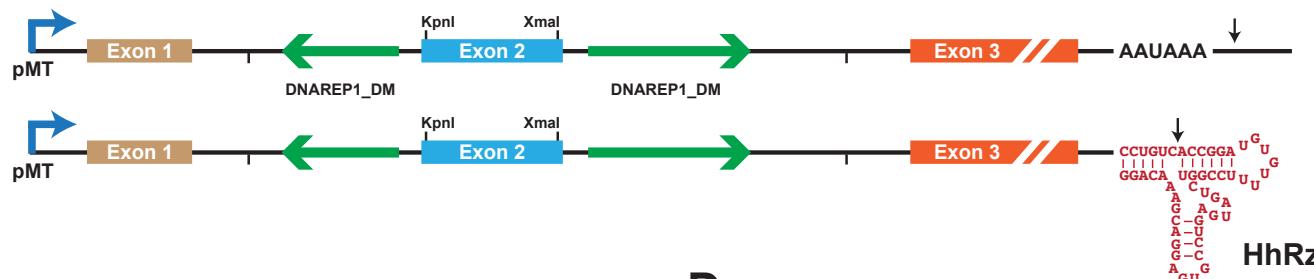
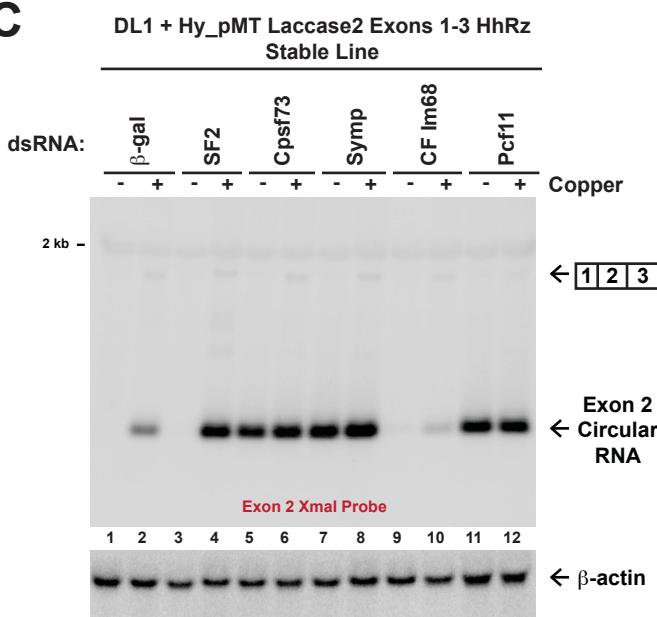
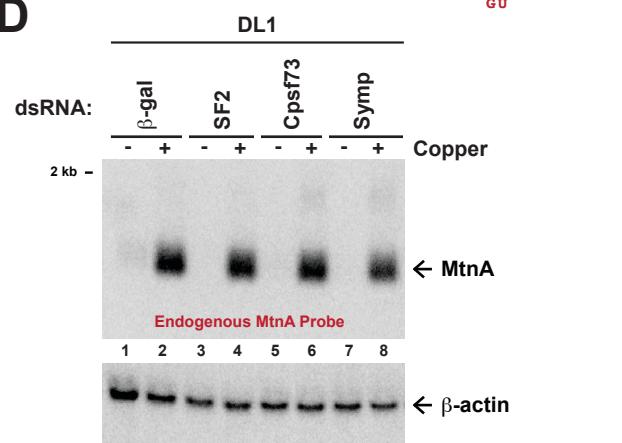
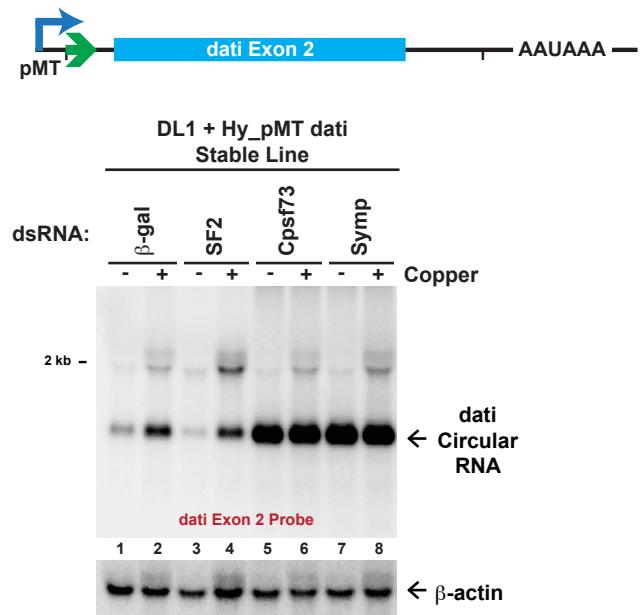
A

B

C

D

E


Table S1. Oligo sequences, Related to STAR Methods

Northern Probe	Sequence	Main Figures	Figures where probe is used	Supplemental Figures
Exon 2 Xmal Probe	GCTGAGCTCCCGGG	1B, 1C, 1D, 2A, 2B, 3C, 5A, 5D	S1B, S1D, S1E, S2C, S4C	
Exon 1 Probe	TGATTITTCGTGAGTTGGCAACA	1B, 2B		
Exon 3 Probe	CCATTCCCCATGGTCTCCAC	1B		
Circle Junction Probe	GGTACCTTTCGTGACTTC	1B		
β -actin (<i>Drosophila</i>)	AGCACAGTTGCCGTACAG	1B, 1C, 1D, 2A, 2B, 3A, 3B, 3C, 4A, 5A, 5D, 5E	S1B, S1D, S1E, S2C, S4C, S4D, S4E	
eGFP Probe	GTCA CGAACT TCCAG CAGGAC	5E		
Laccase2 Exon 2 Probe	GCTTAGATTGAGGATGGAGCTCC	2E		
Hygro Probe	GACATATCAGCGCCCTCTA	5D		
Uex Probe	GGCTTTAGACATGCCGGCC	3A		
PlexA Probe	GGTCCCGGCGCAATAATGC	3B, 4A		
Endogenous MtnA Probe	AAGATGCAAGCCTCTACTC		S4D	
dat Exon 2 Probe	TCATCTCCTGCTGGGCTTG		S4E	

dsRNA Target	Forward Primer	Reverse Primer	DRSC #
β -gal	TAATACGACTCACTATAGGG CTGGCGTAATAGCGGAAGAGG	TAATACGACTCACTATAGGG CATTAAGCGGAGTGGCAACA	---
Cpsf73	TAATACGACTCACTATAGGG AGCTACTGCTGATGCCATT	TAATACGACTCACTATAGGG GTTACCGCTAACCGTCACT	DRSC28954
SF2	TAATACGACTCACTATAGGG CGACAGTGGTGAAG	TAATACGACTCACTATAGGG CGATGGCGAAAGCGAGAT	DRSC16845
SF3b1 (CG2807)	TAATACGACTCACTATAGGG GTCATTTTATGAAAGCTTGAATTGA	TAATACGACTCACTATAGGG TGCCCAAGCTGTAATTTG	DRSC00535
SF3a1 (CG16941)	TAATACGACTCACTATAGGG AGCTGGAGCAGGAGG	TAATACGACTCACTATAGGG TATCCGTAGCGGCTCAG	DRSC15166
SF3b2 (CG3360)	TAATACGACTCACTATAGGG CCCGAAACTTAAACACTGAC	TAATACGACTCACTATAGGG CGACAGCAATGCCAACT	DRSC00519
SF3d3 (CG13900)	TAATACGACTCACTATAGGG GGGAGCTGAACGAGTACAC	TAATACGACTCACTATAGGG GACAGAGGCCAAATGGA	DRSC03370
SF3d4 (spx)	TAATACGACTCACTATAGGG ATACTCGAGCGCAGAGAT	TAATACGACTCACTATAGGG GTGTCACTAGCTCCCATATT	DRSC23542
SF3d5	TAATACGACTCACTATAGGG GGTCAGACGTACAATATCCA	TAATACGACTCACTATAGGG TAATCTCTCATGTTTGTTG	DRSC14460
SF3d6	TAATACGACTCACTATAGGG CTCTGCCCATTTAGAGTGGCC	TAATACGACTCACTATAGGG AAGGTTTGTCAGGCAGG	DRSC26462
Phf5a	TAATACGACTCACTATAGGG GAAAAAGACGATGGAAAGTGT	TAATACGACTCACTATAGGG GCAGAGACTCGAGTAGTAG	DRSC03261
SF3a2 (CG10754)	TAATACGACTCACTATAGGG AAAGAACCTCTGGGCTCT	TAATACGACTCACTATAGGG GGATCTTAAGGTGGCAAATTGGA	DRSC09801
SF3a3 (noi)	TAATACGACTCACTATAGGG CAGAGAAAACGCCGTCGGAC	TAATACGACTCACTATAGGG ACCCTGGTTTTAGATCTT	DRSC26162
snRNP-U1-70K	TAATACGACTCACTATAGGG GAGAGAACCTCTGGGCTA	TAATACGACTCACTATAGGG GTCTGGTACAGGAGACT	DRSC28803
snRNP-U1-C	TAATACGACTCACTATAGGG CAAAGTACTATGGCAGACT	TAATACGACTCACTATAGGG CTGGTGGCGCTTCATGATTC	DRSC15804
CG7564	TAATACGACTCACTATAGGG GGTCTCGAACGCCCTTTCG	TAATACGACTCACTATAGGG GCCAACCTTCTGCAATCTC	DRSC22225
CG3198	TAATACGACTCACTATAGGG ATAACAGGGCCATCACAAAG	TAATACGACTCACTATAGGG TTGGGAGTAGGCCAGAGTGT	DRSC24615
CG1646 (Prp39)	TAATACGACTCACTATAGGG CCCAGCAAACTGGTAAATAA	TAATACGACTCACTATAGGG TGCTCAGATGTTGGGACTG	DRSC25186
CG3542 (Prp40)	TAATACGACTCACTATAGGG CGAGGACTGTATTTAATTAG	TAATACGACTCACTATAGGG GCGCCGTTTACAGGTGTT	DRSC00612
Rox8 (TiAR)	TAATACGACTCACTATAGGG CAAGCCAGCAGGAGAATAG	TAATACGACTCACTATAGGG TGATGGCAGGCTTTTACG	DRSC29536
snf	TAATACGACTCACTATAGGG TACCAACCAAACCGATTACAT	TAATACGACTCACTATAGGG CTTCCTCATGGTACCGG	DRSC18835
Prp8	TAATACGACTCACTATAGGG ATTTGGAAACCTGTACGGAAA	TAATACGACTCACTATAGGG TTAAACTATCGTCCTTTAACC	DRSC07293
obe	TAATACGACTCACTATAGGG GGAAGTGAACCGAACGCTG	TAATACGACTCACTATAGGG GTGGGGCCTAACAGAC	DRSC15736
CG3436 (SNRNP40)	TAATACGACTCACTATAGGG ACGACCAAATAAGCGTAATT	TAATACGACTCACTATAGGG GAAACCGAAAGAACGAAATAC	DRSC00605
CG4849 (EFUD2)	TAATACGACTCACTATAGGG AGCGGACCTCTGCCATAAAC	TAATACGACTCACTATAGGG CCATACAAACGCCGAGTAT	DRSC27175
CG6841 (Prp6)	TAATACGACTCACTATAGGG CAGGCCAACCGTGTGT	TAATACGACTCACTATAGGG GTTGGAGCAGACTCAC	DRSC10724
CG10333 (Prp28)	TAATACGACTCACTATAGGG TTTCAGGAAAATGATCCAC	TAATACGACTCACTATAGGG TGCGGGTTTACAGCTC	DRSC29270
Dim1	TAATACGACTCACTATAGGG AATACCGCTGGTCACTGATT	TAATACGACTCACTATAGGG GTTGGAGTGTAGATGAGTAGT	DRSC00563
Prp3	TAATACGACTCACTATAGGG AACGCCCTTGTCACTTGT	TAATACGACTCACTATAGGG CTTGTGTGGTACGAC	DRSC30866
U4-U6OK (Prp40)	TAATACGACTCACTATAGGG GGACAGCTACAGGAGTGC	TAATACGACTCACTATAGGG TTGGCACTGGATTTCTG	DRSC42502
hoip (SNU13)	TAATACGACTCACTATAGGG GAAGTTAATCCAAAGGATT	TAATACGACTCACTATAGGG AGGAGCTTCGATC	DRSC03546
CG17266 (PP1H)	TAATACGACTCACTATAGGG TGACCAAGCATACGCGAAC	TAATACGACTCACTATAGGG TAATGGTCACTGGGGAGCTT	DRSC24207
U2af38	TAATACGACTCACTATAGGG CTGATTCTGGAGCTGTGAAAC	TAATACGACTCACTATAGGG TAATGGGTTGGGGTTCACTA	DRSC28068
CG3294	TAATACGACTCACTATAGGG ATCTGAACACCCGGTGTAC	TAATACGACTCACTATAGGG ACTCCTACAGCGTGTGACC	DRSC26228
U2af50	TAATACGACTCACTATAGGG CGGCCATTCGCCAACG	TAATACGACTCACTATAGGG AGAACGGCTTAAATGC	DRSC20297
SF1	TAATACGACTCACTATAGGG CGCAGTCAGCTGTGCA	TAATACGACTCACTATAGGG TTGGTACGCTTGTGTGTT	DRSC16844
Cbp20	TAATACGACTCACTATAGGG CCCCTTCCCACCAACA	TAATACGACTCACTATAGGG CTCTTCCTCCATCCACACAA	DRSC16501
Cbp80	TAATACGACTCACTATAGGG CCCCATTTCTGCACTACAC	TAATACGACTCACTATAGGG TACGCCACGCTGCAACACC	DRSC18450
HeL25E (UAP56)	TAATACGACTCACTATAGGG CTGAGCTCTGGAGCTTGCAT	TAATACGACTCACTATAGGG ATGCCGAGCTGTCCTTAAACA	DRSC30679
CG6227 (Prp5)	TAATACGACTCACTATAGGG ATCTGGGTGCTAAATCGGAG	TAATACGACTCACTATAGGG GCACTAGTGAACCGAGGAAT	DRSC27434
Prp31	TAATACGACTCACTATAGGG GGACAGGAAAGCGAAGAA	TAATACGACTCACTATAGGG TTGGGGCAGGCCCTTTA	DRSC10735
Cdc5	TAATACGACTCACTATAGGG GCAGAGTCAGACTCTGTC	TAATACGACTCACTATAGGG TGATGGTCACTAGCTGAATAAAC	DRSC08577
I(2)37Qb (Prp2)	TAATACGACTCACTATAGGG CTACGGGTCAAGGTAGAT	TAATACGACTCACTATAGGG GAGCCCTTACGAGCACC	DRSC24383
pea (Prp22)	TAATACGACTCACTATAGGG ATGGAGCAGCTCGAGAGTT	TAATACGACTCACTATAGGG TTGAGGTGTCATGGGCAA	DRSC29708
II(2)50Qb (Prp16)	TAATACGACTCACTATAGGG GGCGGGAGCACAGGAC	TAATACGACTCACTATAGGG GGCGCACGTGTCGATAAAAC	DRSC19542
CG6015 (CDC40)	TAATACGACTCACTATAGGG ATGCGACTGACTATAGGG	TAATACGACTCACTATAGGG GATCTTCCTGCGAGGTGC	DRSC28104
Slu7	TAATACGACTCACTATAGGG GCGCAAGCTGCTGAAATCCAT	TAATACGACTCACTATAGGG AGCCAGTGGGATTGTCAC	DRSC23582
Prp18	TAATACGACTCACTATAGGG GGCGACCACTTCTATAATT	TAATACGACTCACTATAGGG GTTGGGGTTGGGTGAG	DRSC16802
Rat1	TAATACGACTCACTATAGGG AAAGAGTCGTTACAGAAGCG	TAATACGACTCACTATAGGG GCCCTGATAGTGCCTTAAGAACCC	DRSC02044
Symp	TAATACGACTCACTATAGGG CGAGCTTAATTCCGTTCTGC	TAATACGACTCACTATAGGG TTCCCGAGTACTTTTCGCG	DRSC29255
mts	TAATACGACTCACTATAGGG CGAGCTTACAGGCGACAGT	TAATACGACTCACTATAGGG CACGAGGCCGAGATTCCCC	DRSC03574
Pp2A-29B	TAATACGACTCACTATAGGG CGCAGGAGCTCTTC	TAATACGACTCACTATAGGG GACGATGCCGGCGAGCGA	DRSC03400
ATPsyn β	TAATACGACTCACTATAGGG CGTCTGAGCTTGTGATGAT	TAATACGACTCACTATAGGG TTGGTGCCTTGTGTTTCG	DRSC34212
smt3	TAATACGACTCACTATAGGG GTCTGACAAAGAAAGGGAG	TAATACGACTCACTATAGGG ATGGAGCCGACACAGTC	DRSC03611
CF Im88 (CG7185)	TAATACGACTCACTATAGGG GCCGCCTACAGCTTTATG	TAATACGACTCACTATAGGG ATTCTGGTGGAGGTTCTGAGT	DRSC10781
Pcf11	TAATACGACTCACTATAGGG GCATCTTCCCTTTGGAGA	TAATACGACTCACTATAGGG GTCACGGGGGTATCTTCA	DRSC23350

Human ASOs [Note: "m" represents 2'-O-Methyl (2'OMe) modification]

Scramble ASO	mGmCmUmUmCmCCTGAGGTTmCmUmCmC
MATR3/PAIP2 Intergenic ASO	mTmAmCmTmTmATGCTGCGACmTmGmAmTmC

qRT-PCR Primers	Forward Primer	Reverse Primer	
DROSOPHILA			
β -actin (Act42a)	CTCGGTCACCATGGAAAGATT	TTCGAGATCCACATCTGCTG	
Laccase2 Circular RNA	GCCTCGGAGAAATTCTGACTATCA	ACATGTTGCTGCCAGAACGAC	
Laccase2 mRNA (Ex 2/3 Junction)	GCCTCGGAGAAATTCTGACTATCA	CAACTCCCATCAGCTAACACACAGTA	
Uex Circular RNA	ATTCGATAGTTCGCCGCG	GCAACGGATTTCAGCATCTTACT	
Uex mRNA (Ex 4/5 Junction)	GCACGGATTTCAGCATCTTACT	GTCTGGTCACCTTGCTGTT	
Plexa Circular RNA	ATTTCTTCGTTCTGGTGTCTA	CCCGACATGCCATTGTTCTA	
Plexa mRNA (Ex 2/3 Junction)	CCCGACATGCCATTGTTCTA	GTGTTAACACATGGCCACCGGTAC	
ps Circular RNA	TCGATCTGGAAATTAAGACAA	GACATTCGTTGTAGCTGCT	
mbn Circular RNA	GAGGATCACTGAATTGGCGG	AATTGTCGCGAACGCTTGTGTT	
CG42663 Circular RNA	AGAGCCAAATTGTCGGGG	ATTTGCCCTCAGCTTGGAC	
Ecd4 Circular RNA	TTCAAGCTCTAGTTCCTGGAG	TTTCGTCAGATTGCGCTG	
Cpsf73	AGCAGGGACCATGGGTTAGT	TTCCCGAGTCAGCATTTCT	
SF2	ATTCAGGAGCTTCTGGCAAA	CGCATCTTGCACAACTC	
CG2807 (SF3b1)	GTCATACACCTTCTGGCAAA	CTGGGTTCAGCTCTTAGGAC	
CG16941 (SF3a1)	CGCGGTCGAAATGAGGAATAA	GGGGATCTCCCGGCAATTAGAAAGT	
SF3b2 (CG3605)	GATTCGGGATACAGGGTAAC	CTGGGTTCAGGAAACTCATTA	
SF3b3 (CG13900)	CTTGTGACTACCTCCACGT	CGGGGACTCTCTGTTATCCT	
SF3b4 (spx)	CCAAATGCAACGGGCTATG	CGAATGGGCTTGTGCGGAAAG	
SF3b5	GCAGATTAACATTCTCGGCTTA	TCCAGCTGGCTTGTGGATATT	
SF3b6	ATTTCGACCTTCTGCCAGATAC	TCACAGGGCTTCTTGCAC	
Phf5a	ACTGCAAGTCTGCACCATTC	AACAGGTCACTTGGAGGCTG	
SF3a2 (CG10754)	CATGTTCAATGGCGCTTCATC	TTCTGTGGTAATGCCGTTGC	
SF3a3 (noi)	CGACATGAGTGGCCTAGTGG	CGACGCCAAAGTTGGAGGTACAG	
Rat1	CGCTGTGTCGATTGTCAC	GAGATCCTGCTGATCTCATCTT	
HUMAN			
PAIP2 mRNA	GATCTCCCACAAACTATGGA	GGATTCAGATTGCTTCTGAC	
circPAIP2	CCACAAACATATGGACCAAAT	GATGATGCTTGGGCTAGTC	
MATR3-PAIP2 Intergenic 1	ACTTTGACAGGGTTGGAAATC	GTGATGTTTCGAAATTCTCTAT	
MATR3-PAIP2 Intergenic 2	TGTCTCTCTACATCTGCTCA	CACTGTAAATTACACCTCTCT	
pre-PAIP2	CACAGATTCATTAGCACAGTT	CTGTCCTTCACTACTGGAAAC	

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Related to Figure 1.

Short intronic repeats are sufficient for production of circular RNAs that have long half-lives.

(A) Exon/intron structure of the *D. melanogaster laccase2* locus. To generate the Hy_pMT Laccase2 Exons 1-3 plasmid, exons 1, 2, and 3 with shortened intervening introns were cloned downstream from the copper-inducible metallothionein A promoter (pMT). Inverted *DNAREPI_DM* repeat sequences (green) are present in the introns flanking exon 2. A 514-nt circular RNA is formed when the 5' splice site at the end of exon 2 is joined to the 3' splice site at the beginning of exon 2. The expression plasmid additionally contains ampicillin (Amp) and hygromycin resistance (HygroR) genes. **(B)** DL1 cells were transfected with the Hy_pMT Laccase2 Exons 1-3 plasmid, CuSO₄ added for 14 h, and total RNA isolated. The circular RNA was resistant to RNase R digestion, unlike the linear mRNA or endogenous β-actin mRNA. **(C)** Schematics of Hy_pMT Laccase2 Exons 1-3 expression plasmids containing intronic deletions. **(D)** Deletion plasmids were transfected into DL1 cells, CuSO₄ added for 14 h, and total RNA analyzed by Northern blotting. Most of the upstream *DNAREPI_DM* repeat sequence was dispensable for exon circularization, and pre-mRNAs containing long vs. short upstream repeats circularized at similar frequencies. **(E)** Further deletion constructs revealed that ~100-nt of each repeat (beginning at nt 450 for the upstream repeat and ending at 1245 for the downstream repeat) was sufficient for efficient exon circularization. **(F)** A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1-3 plasmid was treated with CuSO₄ for the indicated amounts of time. Linear mRNA and circular RNA levels were quantified using ImageQuant from four independent Northern blot experiments and were normalized to the circular RNA levels observed at 8 h. Data are shown as mean ± SEM. **(G)** To measure RNA half-lives, the stable cell line was

treated with CuSO₄ for 3 h followed by chelation of the metal with BCS for the indicated amounts of time. Linear mRNA and circular RNA levels were quantified using ImageQuant from three independent Northern blot experiments and were normalized to the RNA levels observed at 0 h BCS treatment. Data are shown as mean ± SD. Cross-hybridization of the Northern probe with ribosomal RNAs is responsible for most of the linear mRNA signal detected after 12 h.

Figure S2. Related to Figures 2 and 3.

Depleting spliceosome components results in increased circular RNA levels.

(A) A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1-3 plasmid was treated with the indicated dsRNAs for 3 d and CuSO₄ was added for the last 14 h. qRT-PCR was then used to verify that each dsRNA efficiently depleted its target gene. Data from two independent experiments were normalized to β-actin (Act42a) and are shown as mean±SD. **(B)** A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1-3 plasmid was treated with the indicated dsRNAs for 3 d and CuSO₄ was added for the last 14 h. 2 h before collection, cells were replated on 96-well dishes, allowed to attach for 2 h, and then fixed. Cells were stained with Hoechst and nuclei quantified by automated microscopy. Data from three independent are shown as mean±SD. **(C)** The RNA samples from **B** were subjected to Northern blot analysis to determine the amounts of linear and circular RNA generated from the Hy_pMT Laccase2 Exons 1-3 reporter. Although depletion of mts, Pp2A-29B, ATPsynβ, or smt3 affected cell viability, these dsRNAs largely did not affect the ratio of linear vs. circular RNA generated from the reporter.

Figure S3. Related to Figure 4.

Depleting spliceosome components results in increased PlexA circular RNA levels.

DL1 cells were treated with the indicated dsRNAs for 3 d. qRT-PCR was then used to quantify expression of the endogenous PlexA linear mRNA and circular RNA. Data from three independent experiments were normalized to β -actin (Act42a) and are shown as mean \pm SD.

Figure S4. Related to Figure 5.

Readthrough transcription from the copia LTR promoter results in production of a circular RNA from the downstream minigene.

(A) A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1-3 plasmid was treated with the indicated dsRNAs for 3 d and CuSO₄ was added for the last 14 h. qRT-PCR was then used to verify that each dsRNA efficiently depleted its target gene. Data from two independent experiments were normalized to β -actin (Act42a) and are shown as mean \pm SD. **(B)** The polyadenylation signal downstream of the *laccase2* minigene was replaced with a self-cleaving hammerhead ribozyme (HhRz) sequence, thereby generating the Hy_pMT Laccase2 Exons 1-3 HhRz plasmid (bottom). **(C)** A DL1 cell line stably expressing the Hy_pMT Laccase2 Exons 1-3 HhRz plasmid was treated with the indicated dsRNAs for 3 d and, where noted, CuSO₄ was added for the last 14 h. Northern blots were used to examine expression of reporter-derived transcripts. Depletion of Cpsf73, Symp, or Pcf11 each caused circular RNA expression to increase and become constitutive. The canonically spliced linear mRNA failed to accumulate in cells, indicating that the HhRz-generated mRNA 3' end is likely unable to prevent decay by 3'-5' exonucleases. **(D)** Northern blots were used to examine endogenous metallothionein A (MtnA) expression in DL1 cells that had been treated with the indicated dsRNAs for 3 days. Where noted, CuSO₄ was added for the last 14 h. **(E)** The *laccase2* minigene from the Hy_pMT Laccase2 Exons 1-3 plasmid was replaced with a 1487-nt region of the *Drosophila* *dati* pre-

mRNA that can generate a 1120-nt circular RNA from exon 2. A DL1 cell line stably expressing the Hy_pMT dati plasmid was generated and treated with the indicated dsRNAs for 3 d. Depletion of Cpsf73 or Symp caused dati circular RNA expression to increase and become constitutive.

Figure S5. Related to Figure 6.

Identification of endogenous readthrough transcription events in human PA1 cells.

(A) Computational pipeline used to detect circular RNAs that may be generated from readthrough transcription events in PA1 cells. CIRCexplorer (Zhang et al., 2014) was used to identify circular RNAs from the population of unmapped reads. By calculating RNA expression levels in the intergenic regions between gene pairs, potential polycistronic transcripts were identified. **(B)** Identification of 39 loci where there is readthrough transcription from the upstream gene and a circular RNA produced from the downstream gene. **(C)** Nascent RNA-seq data highlighting the intergenic region between the *MATR3* and *PAIP2* genes. Intergenic regions targeted by qRT-PCR primers in **Figure 6D** are shown.

SUPPLEMENTAL TABLE LEGENDS

**Table S1. Related to STAR Methods.
Oligo sequences.**

Methods S1. Plasmid Details, Related to Figures 1–3, 5, S1, S2, and S4

All *Drosophila* expression plasmids were generated from the **Hy_pMT EGFP SV40 pA Sense** plasmid (<https://www.addgene.org/69911/>), which is a modified form of the pMK33/pMtHy plasmid. In brief, the metallothionein promoter (marked in blue) drives expression of the EGFP ORF (marked in green) that terminates in the SV40 polyadenylation signal (marked in pink). HygroR (marked in red) is driven by a copia transposon LTR promoter and terminates with an SV40 late poly(A) signal (marked in gray). An Amp selectable marker is also present. The full plasmid sequence is as follows:

```
GATCAATTCTGGTGCAGGACAGGAATGGTGCCCCATGTGACTAGCTTTGCTGCAGGCCGCTCATCCTCTGGTCCGATAAGAGACCCAGAACCTC  
CGGCCGCCACCAGCCACCCCATACATACTGGTACGCAAGTAAGAGTCGCTCGCATGCCCATGTGCCCAACCAAGAGCTTCCATCC  
CATACAAGTCCCAGGGAAACGCAACAACTTCTCGGGCAGAACAAAAGCTCTGCACACGTCCATCGAATTGGAGGCCGGCGT  
GTGCAAAGAGGTTGAACGAAAGACCCGCTGTAAAGCCGCTTCCAAATGTAAACACCGAGAGCATCTGGCAATGTCATCAGTGTG  
GTCAAGCAGCAAATCAAGTGAATCATCTCAGTGAACATAAGGGGGATCTCGAGGTACGATGATAAGCTGATATCACCATGGTGACCAAG  
GGCGAGGAGCTTCACTGGGCTGGTGCCTACTCTGGTCAAGCGTGAAGCTCGAGCTGGCGACGTAAACGCCACAAGTTCAGCGTGTGGCGAGGGCG  
ATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCCGGAAGCTGCCGTGCCCCTGCCCAACCTCGTGAACGCCCTGACCTACGGCGT  
GCAGTCTTCAGCCGTACCCCAGCACATGAAGCAGCAGCTTCTCAAGTCCGCACTGCCCCAAGGCTACGTCCAGGAGCGCACCATTCTTCTC  
AAAGGACGACGCAACTACAAGACCCCGCCGAGGTAAGGTTTGAGGGGACACCCCTGGTAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG  
ACCGCAACATCTGGGCAACAGCTGGTACAACATAACGCCAACACGTATGATTCATGGGCAACAGGCAAGACGGCATCAAGGTGAACCT  
CAAGATGCCACAACTCGAGGAGCCAGCTGGTCACTCGCCGACCCATAACAGGAGAAGCCGATCCAGGGCTGCTGCCGAGGGGGCTGCTGCCG  
AACCAACTACCGACCCCAGTCCCTGAGCAAAGCCCAACGAGAAGCGCAGTACATGGTCTGTTGGAGTTCGACCGCCGAGGATCA  
CTCTGGCATGGAGACTGTACAAGTAAAGCAGGGCAACTTGTATTGAGCTTAAATGGTACAATAAGCAATGATCACAAATTTCACA  
AATAAAGCATTTTCACTGCATTCTAGTGTGGTTGCTCAAACCATCAATGATCTTAACTAGTGGATCCACTAGGGCGCCGACCGAGGC  
TGGATGGCCTTCCCATTATGATTCTCGCTCCGGCGCATGGGATGCCGCGTGCAGGCCATGCTGTCAGGCAAGTAGATGACGCCATC  
AGGGACAGCTTCAAGGATCGCTCGCGCTTACAGCCTAACCTCGATCATGGACCGCTGATCGTCAAGCGCATGGGCTGCTGCCCTCGGGAGC  
ATAGGAACGGGGTGGCATGGATTGAGGCCGCGCTTACCGATTACCTGCTGCTGCCCTTCCGGTGTGCGGTGATGGAGCCGACCTCGAC  
ATAGGAAGCCGGGCACCTCGCTAACGGATTACCCACTCCAAGAAATTGGAGCCAATCAATTCTGGCGGAAGACTGTAATGCGCAAAC  
AACAGTCTGCACCGTGGCAAGCAACATGAATGGTCTCGGTTTCCGTGTTTGTAAGTCTGGAAACCGCGAAGTCAGCGCTTCCGCTCC  
TCAGTACTCGCTCGCTGGTGTGCGGAGCGGTATCGCTACTCAAAGGCGTAAACCGTTATCCACAGAAATCAGGGGATAACGC  
AGGAAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCTTGTGGCTTTCATAGGCTCCGCCCTGACGAG  
CATCACAAAAATCGACGCTAACGTAGAGTGGCGAAACCCGACAGACTAAAGATAACCCAGGGCTTCCCTGGAAACCTCCCGCTGCTCTC  
CTGTTCCGACCTTGGCGCTTACCGGATACCTGTGCTTCCCTGCGGAAAGCTGGCGCTTCTCATAGCTCAGCCTGATGGTATCTCAGTTC  
GGTAGGTCGTTCTGCTCCAAGCTGGCTGTGCAAGCACCCCCGGTCAAGCCGCGCCATTGCGCTTACCGTAACTATGTTAGCTGAGTC  
CCGTTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGACCGAGGTATGAGCGGTCTACAGAGTTCTGAGGG  
GGCCTAACTACCGTACACTAGAAGGACAGTATTGGTATCTGCGCTGTGTAAGCCAGTTACCTCGGAAAAGAGTTGGTAGCTCTGATCC  
CAAACAAACCACCGCTGGTAGCGGTGTTTGTCAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAAGATCTTGTATTTCT  
ACCGGGTCTACGCTACGGTAAAGGATTGGTCAAGCAGATTACGCGCAGAAAAAGGATCTCACCTAGATCCTTAAATTAAA  
AATGAAGTTTAAATCAATCTAAAGTATATGAGTAACCTTTGTCGACAGTTCAACGTTACAGTCTGAGGACCCATCTCAGCGATCTGCT  
ATTCTGCATCATGATGTGCGTACGCCCCGCTGTCGCTGAGATAACTACAGTACGGAGGGTTACCATCTGGCCCAGTGCTGAATGAC  
GACCCACGCTACCGGCTCCAGATTATCAGCAAAACAGCGACGGCAAGGGCGAGCGCAGAAGTGTCTGCAACTTATCCGCTCATCC  
AGTCTATTAAATTGCGGAGCTAGAGTAAGTAGTGTCCAGTTAAGTGTGCGCAACGTTGCTGCACTGCGCATCGTGGTGTACG  
CTCGTGTGTTGGATGGCTTACCGTCAAGGCGAGTTACGATGATCCCCATGTTGCAAAAGCGGTTAGCTCTC  
GGTCTCCGATCGTGTGCAAGTAAGTGGCGCAGTGTATCACTCATGTTATGCGCAGTCATAATTCTTACTGTCATGCCATCGTAA  
GATGCTTTCTGTACTGGTAGACTCAACCAAGTCATTCTGAGAAATAGTGTATGCGGCAAGCAGTTCTGGCCGCAACACGGGATAA  
TACCGCGCACATGCAAGAACCTTTAAAGTCATCATGGAAAAGCTTCTCGGGCGAAAACCTCAAGGATCTTACCGCTGTTGAGATCCAGT  
TCGATGTAACCCACTCGTCCACCAACTGATCTCAGCATCTTACTTCACCCAGCGTGTGGTAGCAAAACAGGAGGCCAA  
AAAAGGAAATAAGGGCAGACGGAAATGTGAATACTCATCTTCAACTTCAATTATTGAGCAATTTCAGGGTTATTGCTCATGAGCG  
ATACATATTGAATGTTAGAAAATAAACAAATAGGGTTCCGGCACATTTCCCGAAAAGTGCCACCTGACGCTTAAGAAACCATTTATC  
ATGACATTAACTATAAAATAGCGTATCACGAGGCCCTTGTCTCAAGAATTCAATTTGTAAGATTTTTTATCAAAGTCAGTT  
TTCACCAATTCCATCAACCGAGCAAGGCAAAGCTTGAATAATGTTATATACATATCAAATCGCTGCTGACTGCGTATTGAT  
GGCCCAAGGATCATATTATGCAATCAGGATTAGGAGCATGGTCAAAATGCGCACAGGAACATGAAAGACGCCGTTATGCGCAATTAAA  
ATTGGGTTAAATGCTGTGGAAGACTGGTGTGGCGCATTTAAACTTACCTGTTACACATCAACTACTATTGTGAGGTTAACGCCAT  
TTGCAATCAGATGAGAACATGGCTTGTGTCATCAGTCAAGTACGCTCTTATCTCTGCGTCTGGTTTCTCAAGGTTACCTGCTGCTTCT  
ACAACTACTGTATTGTTGCTTCAAGTACATTGAAATTCATGGTCTGAGCTACGTTACGTTACATTAAATTGACACCAAGTCGTATT  
GTGAGCTATGAGCTCAGCTAACATTCTTATTCCATGAATAAGCCGGAAAATATGCAATCTGAAAGTAAATTAAGCAAACTTACTT  
TGACTCAATACCATGCACTTGTGTGCTAGGTTACGCAATTGAGGCGATTATCCGATAACCAAGCGATGACTGTTCCGTTCTGATT  
TTGAAATTGGAAATGTACAATAGTTGCTATATGTCAGTACAGTACGCTTACCTGAGCTCTTATCTCTGGTTTCTCAAGGTTAC  
TTTGTGTTTGTGAACTTCGTGTTGAGTACATTCAATTAAAGCCGCGCTTCGTCATGCGTTCTGCGCTTCGAATGCTGATAACGCT  
TGTCTCCAATGCTTGTGCTGAGCACAATTCAATTGATGGTAAATTAAAAGCCTGTTACGATT  
TTGTTCCGGTGTGGTGGTGTGCACTGCTCTGGGAACGGCAAATGGTTAGGATTGGAAACCCCATCATTGTTGGAATATACTTCAACCT
```

ACAAAAGTAACGTTAACACAACACTACTTATATTGATATGAATGGCCACACCTTTATGCCATAAAACATATTGTAAGAGAATACCACTTTTA
TTCCCTCTTCCTCTTGCTGTAAGTAGTCGTGCTGGTGTCCAGTTGAAATACTTAAATATAATCATAAAACTCAAAC
ATAAAACTTGACTATTATTATTAAAGAAAGGAATATAAATTACAACAGGTTATGGACCTGCAGCCAAGCTGGCGCTCGTCCGGG
GGCAATGAGATATGAAAAGCCTGAACTCACCGCGACGCTCTGTCGAGAGTTCTGATCGAAAAGTTGACAGCGTCTCCGACCTGATGCACTCTC
GGAGGGCGAAGAACTCGTGTGTTTCAGCTCGATGTAGGAGGGCGTGATATGTCCTCGGGTAAATACGTCGCCGATGGTTCTACAAAGATCGT
TATGTTATCGGCACATTGCACTCGGCCGCTCCGATTCCGGAAAGTGTGACATTGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCGCC
GTGACAGGGTGTACGTTGCAAGACCTGCCGAAACCGAAGTGTGCCCCGTTCTGACGCCGGTCGCGGAGGCCATGGATGCCATCGCTGCCG
TCTTAGCCAGACGAGCGGGTTCGGCCATTGCGACGGCAAGGAATCGTCATAACACTACATGGCGTATTTCATATGCCGATTGCTGATCCC
GTGATCACTGGCAAACCTGATGGACGACCCGTCAGTGCCTCGCAGGGCTCGATGAGCTGATGCTTGGCCGAGGACTGCCCGAAG
TCCGGCACCTCGGCACCGGATTTGGCTCCAACATGTCCTGACGGACAATGGCGATAACCGGGTATTGACTGGAGCAGGGGAGGCGATTCGG
GGATTCCAATACGAGGTGCCAACATCTCTTGAGGGCGTGGTGGCTTGTATGCTCCGATTGGTCTGACCAACTCTATCAGAGCTTGGTGA
GGCGCAGGGTGTGACGGCAATCGCCGATCCGGACTGTGGCGTACACAAATGCCCGCAGAGCGGCCGAGGCGATCCGGAG
CTGCTAGAAGTACTGCCGATAGTGGAAACCGACGCCAGCACTCTCCGAGGGCAAGGAATACAGTAGATGCCGACCGAACAGAGCTGATT
CGAGAACGCCCTAGCAGCAACTCGCGGAGCGTAGCAAGGCAAATCGAGAGAACGGCTTACGCTTGGTGCACAGTTCTCGTCCACAGTTCGCT
AACGCTCGCTGGCTGGCGGGGGGGGGCGGTGACTGATTAGCTGGTGGTCCACAGGTCAGATTGTCATGCCAACGAC
TCGGGTGATCTGACTGATCCCGAGATTGGAGATGCCCGCCGTGCTGCCGATTGGTGCAGATTTGTGAAGGAACCTTACTTCTGTTGTA
CATATTGGACAAACTACCTACAGAGATTAAAGCTAAGGTAATAATTAAATTAAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTG
TGTATTTAGATTCCAACCTATGAACTGATGAATGGAGCAGTGGGAATGCCCTTAATGAGGAAACCTGTTGCTCAGAAGAAATGCCATCT
AGTGATGATGAGGCTACTGCTGACTCTAACATTCTACTCCCTAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTCCCTCAGAATTGCTAA
GTTTTTGAGTCATGCTGTGTTAGTAATAGAACTCTGCTTGTCTGCTTGTATTTACACCACAAAGGAAAAGCTGCACTGCTATAAGAAAATTAT
GGAAAAATATTCTGTAACCTTATAAGTAGGCATAACAGTTATAATCATAACACTGTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATT
AATAACTATGCTCAAAATTGTAACCTTAGCTTTAATTGTAAGGGGTTATAAGGAATATTGATGTTAGTGCCTGACTAGAGATCATA
ATCAGCCATACCCACATTGTAAGGTTTACTGCTTAAAAAACCTCCACACCTCCCCCTGAAACCTGAAACATAAAATGAATGCAATTGTTG
TTAATTGTTTATTGCAAGCTTATAATGGTTACAAATAAGCAATAGCATCACAAATTTCACAAATAAGCATTGTTACTGCATTCTAGTGTGG
TTTGTCCAAACTCATCAATGTATCTTATCATGTCG

The following eight plasmids were generated by inserting the indicated sequences into the XhoI and NotI sites of **Hy_pMT EGFP SV40 pA Sense**, thereby replacing the EGFP ORF with portions of Laccase2 or dat1.

Hy_pMT Laccase2 Exon 2

(Exon 2 in yellow)

Note: This plasmid was previously named Hy_pMT Laccase2 MCS-Laccase2 608-1097 (Kramer et al. 2015; <https://www.addgene.org/69890/>)

Used in Figure 1B

Used in Figure S1E

Hy-pMT Laccase2 Exons 1-3

(Exon 1 in green, Exon 2 in yellow, Exon 3 in blue)

Used in Figures 1B, 1C, 1D, 2A, 2B, 3C, 5A, 5D

Used in Figures S1B, S1D, S1E, S2C

CTATAACTGGCACGGTATATGGCAACGTGGATCGCAGTATTATGATGGGGTCCCTCGTTACTCAGTCCCCCATTAGCAGGGTAATACCTTCCG
 TTACAGTGGACAGGAACCGCTGGACACATTTCGGCATGCTCATACGGCTACAGAAACTAGTGGACTCTACGGCAGCGTCGTGGTTCGACAG
 CCACCATCAAGGGACCTAACCCCCATTTGGCATGCTCATACGGCTACAGAAACTAGTGGACTCTACGGCAGCGTCGTGGTTCGACAG
 GATATCCTGGACGCTTGCGTAAACACAGGCAAGATCGGAGTCATGTTGATTAACGAAAAGGCCAGTTCGTGAACGGACACTGAGTTAT
 GACTAATACTCCGGAAATATTACAATTAGGCCAGGCACGTCCGTTACCGCTTCAAGTATAATGCATTGCTTCAGTCAGCAGCACAGGTA
 ACAATTGAGGGTACGGCATGACGGTTATAGCAACCAGGGTACGGTGAACGGCATCTGTTGATGTTAATACAAAT

Hy_pMT Laccase2 Exons 1-3 Δ300

(Exon 1 in green, Exon 2 in yellow, Exon 3 in blue)

Used in Figure S1D

CTTCAGTTGGCTTAAACCGCTTGGCTCGTACTTCATAAGCGTAATCAATTAGGTGGTGTGCTATTGCTTAGTACACATACTTTAAAAATAT
 ATAGTCTTATGGTGAAACATCCACCCAAAGTATTGTTAAATAGCTAAAGGAATTAAATAAAAAATACAAGAAGAAATAGTGAAGAAATAAAATATACG
 CCATTCAACAACTATGAAAGCCATGGATCTAACGGGTTACCAAACCTTTATTCGGTCTGATATGGGGTGCATTTCACCAATTTCGAAAGT
 ATTGAGTCAGTGGTggcaactcacagaaaaatcaACCTAAAGAACGTAAGGAAATCCCATCTGGTAATTTTTACCTAAGAAGTCTTGAAGT
 ATCATTTGTGATTTGCGACAGTACATATTCCCTTCACTAAATGATTATTAAAGTCAACTACGATATACTCGTATACCTAAACATGACTATA
 TTCCAAAGGGGTGTTCCAAAGATTGGagggtttggcaaatcgtagaaattatacaagacttataaaaattatgaaaaaatacaacaaaattttaaa
 cacgtggcggtacaggtttggcggttagggcgttagagttagggcaggacagggttatcatcgactaggcttgatctgtacagaatatataat
 actttataccgtcccttcatgttacctatTTTcaacgaatctgtataccctttactgtacgattatgggtataATAAGCTAAATCG
 AGACTAAAGtttttgttatatatTTTttttttttttatGCAGAAAGTACCAAAATCAACCAAGATCTACACCAATGGGGCCTTCTGGCAGCAA
 CATGTGGCCTCTTAGTCTCAACTCATTGATTCACCTTTCGGCAACACAGCCGCTTGTGGCAGACTCACCCGCTATAGGTGGAAAGCC
 ACTCTGGGCCGGCCTCTCAATGGCAATAGCCCTGGGCTGGAGCCTACCTCAACCTAGCTGGTAAAGGCGCCCTAACCGCTGGAAAGCG
 TTACCTAAACTCAGCGGGAGTCACCGACCTCTGAGCTTACCACTAAAGTACAGGAACCGCTGTTGAACCCAACCCAAATCTCATTC
 CGCCACTTAGATTTCCTACTAGTCAACTGCCAGCTGGCGCAATCCTGCCGTATCTGCTCTGATGAATGTCAACGCCCTTGCGCGAAGGCG
 AGCCTCCGAGAATTGACTATCATTACGGTGAATACACAGTATTGGGAGCCCGGGAGCTCAGCAGTAAGTATTCAAATTCCAAAATT
 TTTACTAGAAATTACGATTTAAATAGGCAGTTTACTATTGTATAGTTGtagattcgtgaaaagtatgttaacaggaagaataaagca
 tttccgaccatgttataaggatataattcttaaaggatcaatagccaggctgcattccgcattgtccgtctgttatttttattaccggc
 gagaatccatcgaaaactttttccatggctccatggcattttttccatggcattttttcaattgttataattcgtgtaaattttcatcaatt
 tgccaaatccatcgatcggccatggccataacgccttcgttgcattccactatgtgtccatggccatccatggccatccatggcatt
 ccccaatccatcgatcggccatggccatggccatccatggccatccatggccatccatggccatccatggccatccatggccatccat
 gactTATTTCATATACAATGAAAATGAATTAAATCATATGAATATCGATTAGCTTTATTAAATATGAATATTATTGGGCTTAAGGTGTA
 ACCTcctcgacataagactcacatggcgccaggacatggccatggccatccatggccatccatggccatccatggccatccatggccat
 TTCAATTATTGCCAACATTGGAGACACAATTAGTCAGTGGCACCTCAGGCCGCGCAGAAACCTGgtccatggccatccatggccat
 atggccatccatggccatccatggccatccatggccatccatggccatccatggccatccatggccatccatggccatccatggccat
 AGTAAATTTAAATCAACATGAAATTAAATTAAATTTTGACCAATTGCGTACACTTAAATATGCTTTCTTACAATTGAGCA
 AAAGCAACGAATCGTTGGAGCCATTGTCAGTGTCTTAGCTGATGGAGCTTGAACGTGAAACTCTAATCGCAACTGTTGATACCCGG
 AGCATTCAGTGTGAAACGACAAAGTCGTTATGATGTGAGAACCATATGGAGGAATGGAGTAACATACGACGGTATATGCCAAC
 GTGGATCGCAGTATTATGATGGGTGCCCTCGTTACTCAGTCCCCATTCACTAGCAGGGTAATACCTCGTTATCAGTGGACAGGAAACGCTGG
 ACATTTGGCATGCTCACCGCTACAGAAACTAGATGGACTCTACGGCAGCGCTGGTGCACGCCCATCAAGGGACCTAACTCCCAT
 TTATGATTTGATTTAACACACATATAATGCTTACAGTGGTGGTGCATGAAGATGCAGCAGAGAGATATCTGGACGCTTGGCTGTTAAC
 CAGGCCAACATCGGAGTCATGGTAAACGGAAAGGCCAGTTCTGGACCCGAACACTGGATTATGACTAAACTCCGGAAATTTCAC
 AATTACGCCAGGACGCTGTACCGCTTCAAGTAAATGCAATTGCTTCAGTCGTGAGCCACAGGTAAACAATTGGGGTACGGCATGCGGTT
 ATAACCAACCGATGGTGAGCCCTGTCATCTGTTGATGTAATAACAAAT

Hy_pMT Laccase2 Exons 1-3 Δ400

(Exon 1 in green, Exon 2 in yellow, Exon 3 in blue)

Used in Figure S1D

CTTCAGTTGGCTTAAACCGCTTGGCTCGTACTTCATAAGCGTAATCAATTAGGTGGTGTGCTATTGCTTAGTACACATACTTTAAAAATAT
 ATAGTCTTATGGTGAAACATCCACCCAAAGTATTGTTAAATAGCTAAAGGAATTAAATAAAAAATACAAGAAGAAATAGTGAAGAAATAAAATATACG
 CCATTCAACAACTATGAAAGCCATGGATCTAACGGGTTACCAAACCTTTATTCGGTCTGATATGGGGTGCATTTCACCAATTTCGAAAGT
 ATTGAGTCAGTGGTggcaactcacagaaaaatcaACCTAAAGAACGTAAGGAATCCCATCTGGTAATTTTTACCTAAGAAGTCTTGAAGT
 ATCATTTGTGATTTGCGACAGTACATATTCCCTTCACTAAATGATTATTAAAGTCAACTACGATATACTCGTATACCTAAACATGACTATA
 TTCCAAAGGGGTGTTCCAAAGATTGGagggtttggcgttagagttagggcaggacagggttatcatcgactaggcttgatccatgtac
 gacttataccgtcccttcatgttacctatTTTcaacgaatctgtatccatTTTTTactgtacgatttttttgttataATAAGCTAAATTGAG
 CTAAGtttttgttatatatTTTTTTGCAAGAAAGGTACCAAATCAACCAAGATCTACACCAATGGGGCCTTCTGGCAGCAACAT
 GTGGCCCACTCTAGTAGTCACCCATTGCTTACCCATTTCGGCAACACAGCCCTTGTGCAGACTCACCGCTATAGGTGGAAAGCC
 CTGGGCCGCTCCATTAAATGCCAACAGCTGGGCTGGAGCTCCATTCTCAACCTAGCTGGTACTAACGCCCTAACCGCTGGAAGCG
 CCTAAACTCACCGGGAGTCACCGACCTCTGCACTGTTATCCAGCTAAATAGATCACAGGAACCGCTGTTGAACCCAACCCAAATCTCC
 CACTTAGTTTCCTACTAGTGCACCTGCCGAGCTGGCGCAATCTGCCGTATCTGCTCTGATGAATGTGCACCGCGTTGCGGAAGGCC
 CTCGGAGAATTGCTACTATCATTTCAGTGGAAATACAGTATTGGGAGCCCCGGGAGCTCAGCAGTAAGTATTCAAATTCCAAAATT
 TAATGAGAAATTGCAATTAGGCAGTTTACTATTGTGatTTGtagattcgtgaaaagtatgttaacaggaagaataaaggat
 ccgaccatgttataaggatcaatagccaggctgcattccgcattgtccgtctgttatttttattaccggccagacat
 cagaactataaaagcttagaaggatgttttagcatcagattctgttagacaaggacgcaggccaatTTTGTGttaatttcat
 ttctcaattggccaaaactggccatggccatTTTGTGactttttttccatattgttataactgtgttaattttcatcaatt
 caaaaaaaaacttttgtcaggccatggccatggccatggccatggccatggccatggccatggccatggccatggcc
 caatcatcgatcccccattggccatggccatggccatggccatggccatggccatggccatggccatggcc
 tTATTTTTATACAAATGAAATTAAATCATATGAATATCGATTAGCTTTATTAAATATGAATATTATTGGGCTTAAGGTGAACC
 TcccgacataagactcacatggcgccaggcacttttttttttttttttttccatTTGTCGGGCTCGCACCCCTCAGCAGCACCT
 AATTATTGCCAACATTGGAGACACAATTAGTCAGTGGCACCTCAGGCCGGCCAGAAAACCTGGTACCGCCCTTCACT
 gacacacaaaacactgttataaacttttttttttttttccatTTGTCAGCTTAATATGCTTTCTTACAATTGAGCAAGT

TTAaaaataaaatcaaaatataatTTTATCGGATTGACCTAATTATAATTATAATTACACAGTCGGTGTCAAGTATGACTCCAAA
CGCAACGAATACGGTTGGAGCCATTGTCAGTGTCTTAGCTGATGGAGTTGAACTGTAATTCTAACCTGCAATCGTATGATACCCGGGCAAGC
ATTCAAGTGTGAAACGCAAAGTCGTTATTGATGTGGAGAACCATATGGAGGGAAATGGAAGTAACTATACACTGGCACGGTATATGGCAACGTG
GATCGCAGTATTATGATGGGTGCCCTTGCTACTCAGTGCCTTACCTCAGCAGGGTAATACCTCCGTATCAGTGGACAGGAAACCGCTGGACACA
TTTTGGCATGCTCATACCGGTCTACAGAAACTAGATGGACTCTACGGCAGCGTGGTCAGCAGGCCACCATCAAGGACCCCTAACCTCCATTAA
TATGATTTGATTTAACACACATATAATGCTTACAGTGGTTGATGAGATGCGACAGAGAGATATCCTGGACGCTTGGCTTAAACACAG
GCCAAGATCCGGAGTCATGTTGATTACGGAAAAGGCCAGTTCGTGCAGCCAAACACTGGATTATGACTAAATCTCGTTGGAAATTACACAAAT
TACGCCAGGACGTCGTTACCGCTTCAATGATAAAATGCATTGCTTCAGTGTCCAGCACAGGTAACAATTGAGGGTCACGGCATGACGGTTATA
GCAACCGATGGTGGAGCCTGTCATCCTGTTGATGTAATACAAT

Hy_pMT Laccase2 Exons 1-3 Δ450

(Exon 1 in green, Exon 2 in yellow, Exon 3 in blue)

Used in Figure S1D

Hy_pMT Laccase2 Exons 1-3 Δ500

(Exon 1 in green, Exon 2 in yellow, Exon 3 in blue)

Used in Figure S1D

CTTCAGTTGGTGTAAACCGCGTGCCTGACTTCATTAAAGCCGTAATCAATTAGGGTGTTGCTATTGCTTAGTACACATACTTAAAATAT
ATAGTCTTATGTTGTGAAACATCCACCCAAAGTATTGTTAAATAGCTTAAAGGAATTAAAAAATACAAGAAGAATAGTGAAGAATAAAATTATACC
CCATTCATACAACTATGAAAGCCATGGATCTTACCGGATTTACCAAACCTTATTTCGTCCTGATATGGGGTCATTTCACCTTCCAAATTCTGAGT
ATTTCGAGTACGTTgttggccaactcaacagaaaaatcaACCTAAGGAAAGCTAAGGAAATCCCATTGTTAATTTCACCTAAGAAGTCTGAAAT
ATCATTGTTGTGATTTGCGAGCTACATATTCTTCAATTAAAGTATTGTTAAGCTATACGTATACTCTGATACATCTAAAGCTGACTATA
TTTCAAGGGGTGTTCAAAGATTGGGttttcaacgaattctgatataccctttactgtacgatattggatataAATAAAGCTAAATCGAGACT
AAAGttttttgttatatatatttttttttatGCGAAAGGTACCAAAATCAACCGAGATTACACCAATGGGGCTCTGCGCAACATG
GCCCACTCTAGTAGTTCAACTCCATTGATTCAACCATTTTCGGCAACACACGGCCTGTGCAGACTCACCCGCTATAGGTGAAAGCCACTCT
GGGCCCCTCCATTAAATGGCAATAGCCTGGGTCTGGAGCTCCATCCTAACCTCTAGCTGGTACTAAGCCAGGCCAACCGCTGGAAAGCGGTTAC
TAAACTCACCGGGAGCTTACCGACCTCTGCACGTTACCTAGCTAATAAGATCACAGGAACCGTCGTGAAACCCAACCCAAATCTCATTCCGCCA
CTTAGATTCTCTACTAGTGCACACTCCGAGCTGGGGCAATTCTGCGCTATCTGCTCTGATGAATGTGCAACCGCCTGTCGCGAAGGCGAGCT
CCGAGAATTGTCGACTATCATTTCAGTGGAAACTACACAGTATTGGGAGCCCCGGGAGCTCAGCTGAAGTATTCAAAATTCAAATTTTA
CTAGAAATATTGATTTTAATAGGCAGTTCTACTATTGTTAGTACTATTGtagattcgtggaaaatgttaacaggaaataaagcattttcc
gaccatgtttaaggatatatattcttaaaatggatcatagccgagtgcattccgcattgtcgctgtcttatttttattaccgcggagacata
gaaactataaaagctagaaggatgagttagcatacagattcttagagacaaggacgcagacaaatgttgcattgtgcacgccttaact
tctcaaatgccccaaactgcgcattccccacatttttgcatttttgcatttttgcataattgttattactcgtgtaaatttccatcaatttgcca
aaaaacttttgcacgcgttaacgcctaaagcccaatttgcgtcagccacactattgaacaattatcaaattttctcattttattcccc
atatctatcgatatccccgattatgaattttatcgccgttcgcattcacactgcgtgatgcgttgcgttgcgttgcgttgcgttgcgttgcgt
ATTTTTATATACATGAAATGAAATTAAATCATATGAAATATGATTAGCTTTTATTAAATATGAAATTATTGTTGGCTTAAGGTGTAACCTC
ctcgacataaagactcacatggcgcggccacattttgcattttgcattttgcattttgcattttgcattttgcattttgcattttgcattttgcatt
TTATTGCAACATGGAGACACAAATTAGTCTGTGCGCACCTCAGCGGGCCGccagaaaaacttgcgcgcgcgcgcgcgcgcgcgcgcgcgc
cacacacacactgtttaaaatatttttgcattttgcattttgcattttgcattttgcattttgcattttgcattttgcattttgcattttgcatt
TGGCTACACTTAATATGCTTCTTACATGAGCAAGTT

AaaaataaataaataatcaaaatataatTTTATCGGATTGACCTAATTATAATTATAATTACACAGT TCGGTGTCAGTATGTACTCCAAACG
 CAACGAATACGGTTGGAGCATTGTCAGTGTCTTAAGCTGATGGAGTGAACGTGAAATTCTAACCGAATCGTATGATACCCGGCAACGCAT
 TCAAGTGTGAAAACGACAAGTCGTTAGTGTGGAGAACCATATGGAGGGAAATGGAAGTAACATACCTGGCACGGTATATGGCAACGTGGA
 TCGCAGTATTATGATGGGGTGCCTCTGTTACTCAGTCCCCCATTCAACAGGGTAATACCTCCGTTATCAGTGGACAGGAAACGCTGGCACATT
 TTTGGCATGCTCATACCGGCTCACAGAAACTAGATGGACTCTACGGCAGCGTCGGTTCGACAGGCCACCATCAAGGGACCTAACCTCCATTATA
 TGATTTGATTAAACAACACATAATGCTTATCAGTGATTGGTGCATGAAGATGCAGCAGAGAGATACCTGGACGCTGGCTGTTAACACAGGC
 CAAGATCCGGAGTCTATGTTGATTAACGAAAAGGCGAGTTCGTTGACCCGAACACTGGATTATGACTAATACTCGTTGAAATATTACAATTA
 CGCCAGGACGTCGTTACCGCTTCGAATGATAATGCTTCAGTCTGTCAGCACAGTAACAATTGAGGGTACGGCATGACGGTTAGC
 AACCGATGGTGACGCTGTGCATCTGTTGATGTAATACAAT

Hy_pMT Laccase2 Exons 1-3 450-1245 (Exon 1 in green, Exon 2 in yellow, Exon 3 in blue) Used in Figure S1E

CTTCAGTTGGCTTAAACGCGTTGCTCGTACTTCATTAAGCGTAATCAATTCAAGGGTGTGCTATTGCTTAGTACACATACTTTAAAAATAT
 ATAGTCTTATGTTGAAACATCCACCAAAGTATTGTTAAATAGCTAACAGGAAATTAAATAAAATACAAGAAGAATAGTAAAATAAAATTACG
 CCATTCAACAACTATGAAAGCCATGGATCTAACGGATTACCAAACATTATTCGTTCTGATATGGGGTGCATTTCATTTCGAAAGT
 ATTGCACTACAGTACGttgcactacagcaaatacAcctaAGAAAGTAAAGGAAATCCCCTGTAATTGGTAAATTTCACCTAAAGAAGTCTGAAAT
 ATCATTTGTGATTGTCGACAGTACATATTCCTTCTTAAATGATTATTAAGTGAATACGTATAACTCGTATACTAAACATGACTATA
 TTCCAAGGGGTGTTCCAAAGGATTGTTGgtacaagatataatatttaccgttcctcatgttacactttcaacgaatctgttata
 cttttactgtacgattatggtaataATAAGCTAACATCGAGACTAAGGTTtattgttatatatatttttttttattttatGAGAAAGTAC
 AAAATCAACCAGATCTTACACCAATGGGGCTCTTGTGCGACAACTGTGCCACTCTAGTTCAACTCCATTGATTCAACATT
 AACACACGGCCTTGTGCGACACTCACCGCTCTAGTTGGAAGGCCACTCTGGGCCGCTCATTAAATGCAATAGCCTGGTCTGGAGCTCCATCC
 TCAATCTCTAGCTGGTACTAACGCCAACCGCTGGAAGCGGTACCTAAACTCACCGGGAGTCTACCGACCTCTGCACGTTATCCAGCTAATA
 AGATCACAGAACGGCTGTGGAACCCAACCCAAATCTCATTCCGCACCTAGATTCTACTAGTGCACACTGCCAGCTGGCGCAATTGC
 GCTATCTGCTCTGATGAATGTGACCGCCTGTGCGAAGGCCTCGAGAATTGCTACTATCATTTACGTTGAAACTACACAGTATTG
 GGAGCCCCGGGAGCCTGAGCTAACATTTTACTAGAAATATTGATTTTACTAGGGCAGTTCTATACTATTGTTATAC
 TATTGtagattcgttggaaaatgttaaaaaatcccgatccatcacccatccaaatccatccccatccatccccatccatccccatcc
 tcaaaaaaaaaaaaaatccatccccatccatccccatccatccccatccatccccatccatccccatccatccccatccatccccatcc
 CTAAATATGTCCTTCTTAAATTGAGCAAGTTTAAaaaataaaatccatccccatccatccccatccatccccatccatccccatcc
 ACACACAGTGCGTCAAGTATGTACTCCAAACGCAACGAATACGGTTGGAGCCATTGTCAGTGTCTTAGCTGATGGAGTTGAACTGGAATT
 TAATGCGAATCGTATGATACCCGGCCAAGCATTCAAGTGTGAAAACGACAAGTCGTTATGATGGAGAACCATATGGAGGAATGGAAGT
 AACTATACACTGGCACGGTATATGCAACGTGGATCGCAGTATTATGATGGGCTCGTTACTCAGTCCCCATTACGAGGGTAATACCTC
 CGTTATCAGTGGACAGGAAACGCTGGACACATTTGGCATTGCTCATACCGGTACAGAAACTAGATGGACTCTACGGCAGCGTCGTGGTC
 AGGCCACATCAAGGACCTAACTCCCATTTATGATTAAACAACACATATACTGTTACAGTGTGGTGCATGAGATGCAGCAGAG
 GAGATATCCTGGAGCCTGGCTGTTAACACAGGCAAGATCGGAGTCTATGTTGAAACGGGAGTCTCGTGAACCGAACACTGGATT
 ATGACTAATACTCCGTTGGAAATATTACACATTACGCCAGGACGTCGTTACCGCTTCGAATGATAAAATGATTGCTTCAGTCTG
 TAACAATTGAGGGTCACGGCATGACGGTTAGCAACCGATGGTACGACCCGTGTCATCTGTTGATGTAATACAAT

Hy_pMT dati Used in Figure S4E

(Exon 2 in yellow)

ttaaatctatatggaaaaacatattgaccacgcccactctgaccctaaaaccggccaaTTTTAAAAATTCAAAAAACCTTTACGCTAACAC
 TTTTTTGTCTTACAGGCAATAAAATGATATGAAATTACGATGCTTAAACGAGGCCAAACTTCGCACAGTGGCACCCAGATCCGTGTCAGGGT
 CGTAAGCAGTTGTCTCCGATGCCAAAAGGCCAGGCCAGCGAGCTGGCAAACTGTATTGGCCATCTCTAACAAATTATCTGGAAATGTT
 AAAATGGACTCCGCAGACTCTGGCAGAACGCCTGGCTGCTCCGGTTCGAAACTTGCGCTTACCAAAACTCTCCCGAATCACAGCAA
 TATCCACACGGCCTGCACCAACTCGTCCCGCAGGAGAACGAGCAGTGGCGGAGTTCAACCAGAAAGGTCCAGCAGCAGGATCAGCGTACTGG
 CAACAAACAGTTCCATGGCAATTCCGGACCAATTCAACTGTGGGCAGCAAGGGGAGTCCAGTCACATTCCGCAGCAGCACTCAGAGCAGTC
 TATGGAAAGCAGCTTCCAGGAGAACGAGCTCGCCCCAGCTGAGTAGCCACCACCTGTTAACGCCGTGCGCCGCTGCGGCGTCCACC
 TCAAGTCCACGGCAGTCAGAATAACCTCTGCCAATAGGCAGTCAGGTTGCAAAATAACCTTCGCAACTACGGACAAGGATCCTAAACGCCCTATG
 TGAGGATCAAGCCAAGCAGGGAGATGGATGCGAAGACGCCCTACAGCGGTTGGAGCTGTCCCACCCCTCTAGTTCAAGGCCAGGCCCAGTCC
 TAGCGCGGAGAGACTACAGCACGTGGCGACCCACAGGCTAGAGAAAATCTCGGAAGGTCAGCAGTATACTGGAACCTGGAGCACCCAGTACGGT
 CCACCGACGATGCACAGTCCAGTGGCCCTCTCATCAGCTGGCAGGGAGCATTCGAGACTCATCAGTGCACGCCAATGTCCTGGAAAC
 AGTTGGTAGCTGAATTGGCGCAGGAGCAAGATGCCACCCACAAATAAGCAAGACCTTTCGCAAGGGTACAGTCCCTCAGCACTCTACA
 ACCCCATCCGGAGGCAGCAGCACGCCACATCAAAATAACGAAAATGCCAATGAAATTCAAGTATTGTTGAAGATTTCTGCTTAT
 CTATATACTTTAAAGGGAACCTGGTATCGAATATCGATGCACTAACAGAGTCTAGTAAAAAAATCAACAGAGAACGCTATTGAGTTAGAC
 TAGACTATATGTAACCCGATATTTAGTGTGTTACGAGCGCAACAAATGAACTTCAAAATTGTTGACTGTTGGACGTTGAGGTAAAGTAGG
 GGGTAGGCAAAAGTGTGTTGGTATGTCATAA

The following plasmid was generated by inserting the indicated sequence into the NsiI and BamHI sites of **Hy_pMT Laccase2 Exons 1-3**, thereby replacing the last 100 nt of exon 3 and the SV40 pA signal with a hammerhead ribozyme (HhRz) sequence.

Hy_pMT Laccase2 Exons 1-3 HhRz

Used in Figure S4C

CCTGTCACCGGATGTGTTCCGGTCTGATGAGTCGTGAGGACGAAACAGG

The following plasmid was generated by inserting the indicated sequence into the NotI and SpeI sites of **Hy_pMT EGFP SV40 pA Sense**, thereby replacing the SV40 pA signal with an alternative 3' end processing sequence.

Hy_pMT eGFP-RVFV NSs HhRz

Used in Figure 5E

aggtaaggctccccccccccccccatccgaccgtAACCCAACTCCCCCTCCCCCAACCCCTggCCTGTCACCGGATGTGTTCCGGTCTGAGTCGTGAGGACGAAACAGG