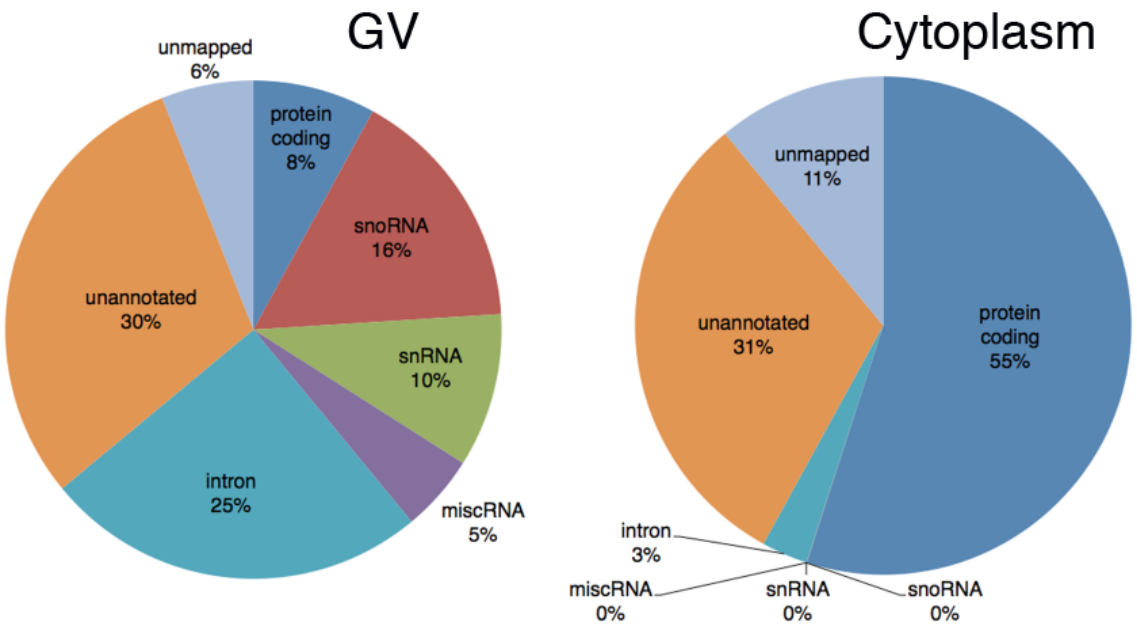
SUPPLEMENTAL MATERIAL

Lariat intronic RNAs in the cytoplasm of *Xenopus tropicalis* oocytes

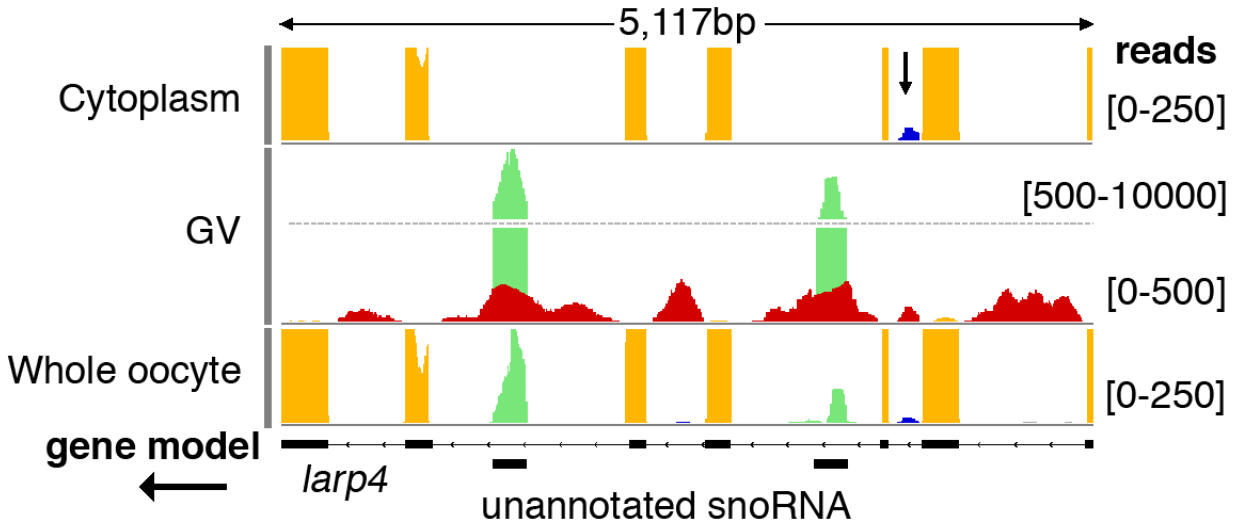
Gaëlle J.S. Talhouarne and Joseph G. Gall

Department of Embryology, Carnegie Institution for Science, Baltimore, MD 21218 and
Department of Biology, Mudd Hall, Johns Hopkins University, Baltimore, MD 21218

Corresponding author:
Joseph G. Gall - gall@ciwemb.edu



Supplemental Figure S1. Relative numbers of reads from an experiment in which RNA samples from *X. tropicalis* oocyte nuclei (GV) and cytoplasm were subjected to deep sequencing. In the nucleus snRNAs, snoRNAs, and intronic sequences make up the majority of annotated sequences. In the cytoplasm, mRNA sequences constitute the major class of annotated sequences, with intronic sequences next in abundance. Noteworthy is the complete absence of snoRNA and snRNA sequences from the cytoplasm, attesting to the lack of nuclear contamination in this fraction.



Supplemental Figure S2. Use of a snoRNA as an internal standard to compare relative abundance of cytoplasmic and nuclear sisRNAs. snoRNA sequences (green) are strictly nuclear, as shown by their absence from a sample of cytoplasmic RNA (top track). They are sufficiently abundant to appear in the whole oocyte sample (bottom track) and they are by far the most abundant sequences in the GV sample (middle track, note the change in scale to accommodate the extremely abundant snoRNA). Thus snoRNA abundance (green) can be compared to cytoplasmic sisRNA abundance (blue) in the whole oocyte sample and to nuclear sisRNA abundance (red) in the GV. From these ratios the relative abundance of cytoplasmic sisRNA to nuclear sisRNA can be approximated, in this case

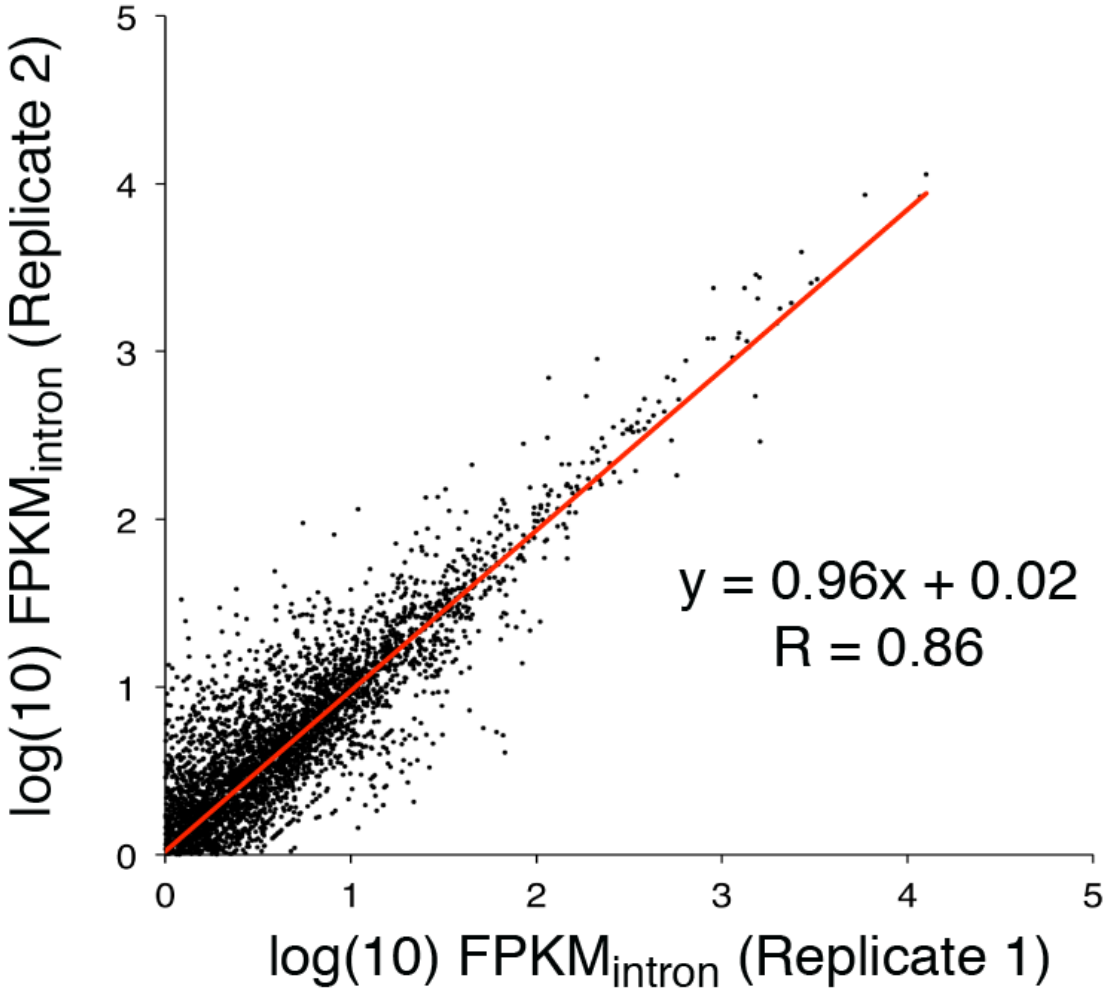
18.9/128.6 X 5604/80.9 = 10.2

18.9 reads/100bp = sisRNA in cytoplasm (blue in whole oocyte)

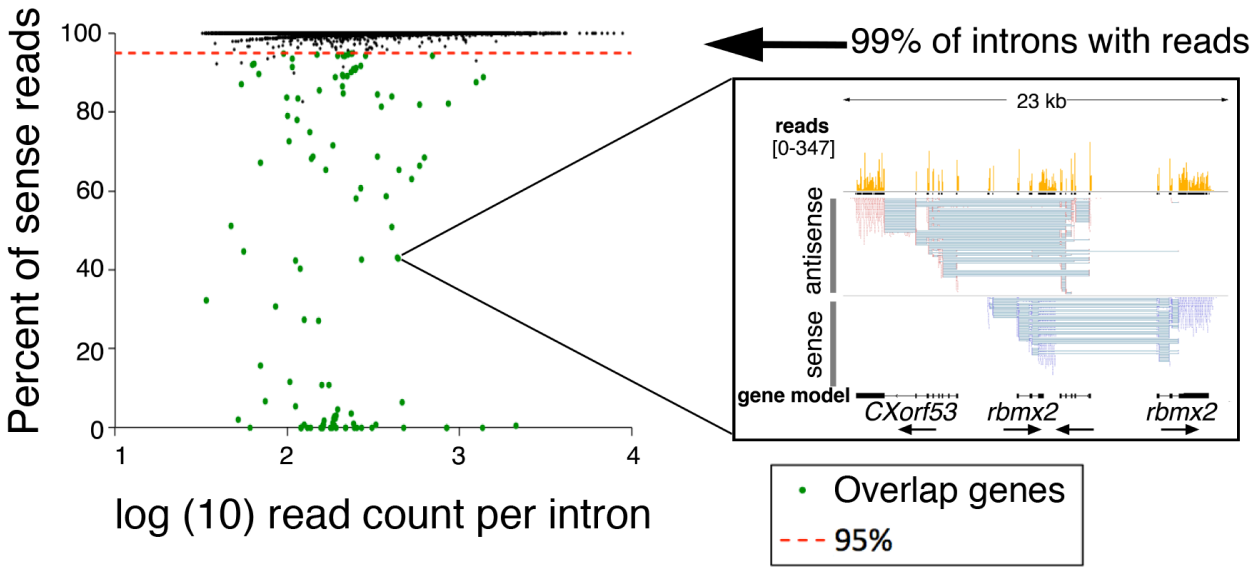
128.6 reads/100bp = average of the 2 snoRNAs (green in whole oocyte)

5603.9 reads/100bp = average of the 2 snoRNAs (green in GV)

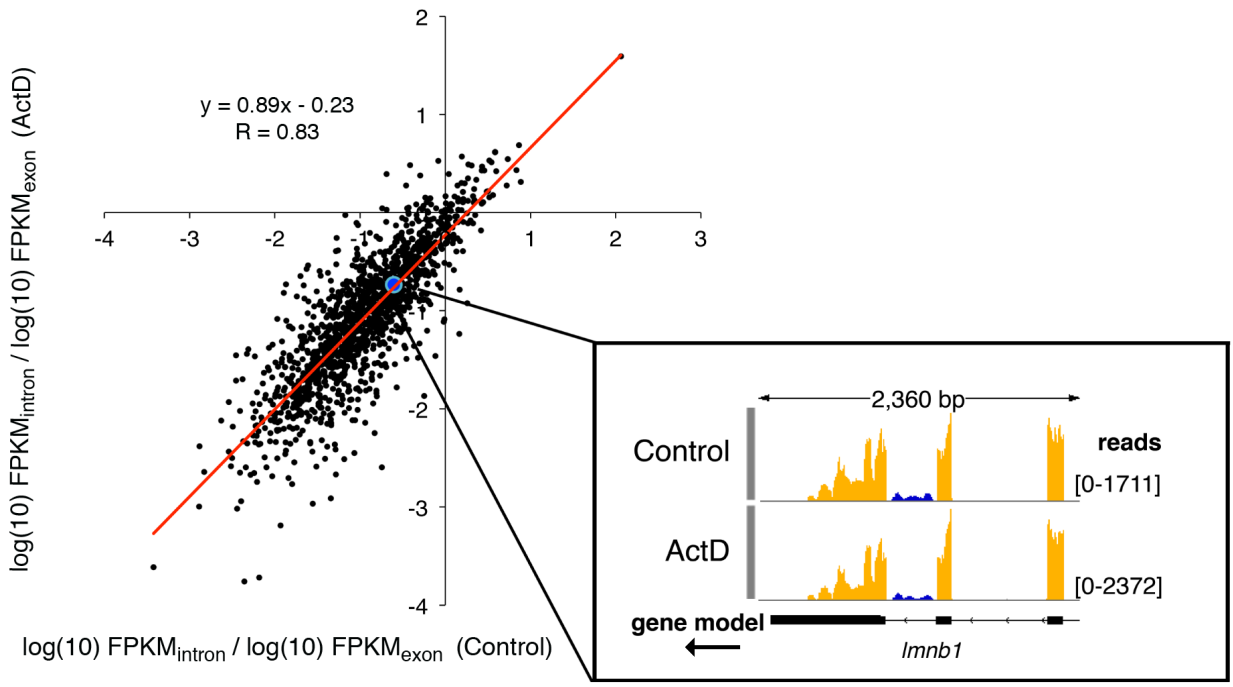
80.9 reads/100bp = average of the nuclear sisRNAs (red in GV)



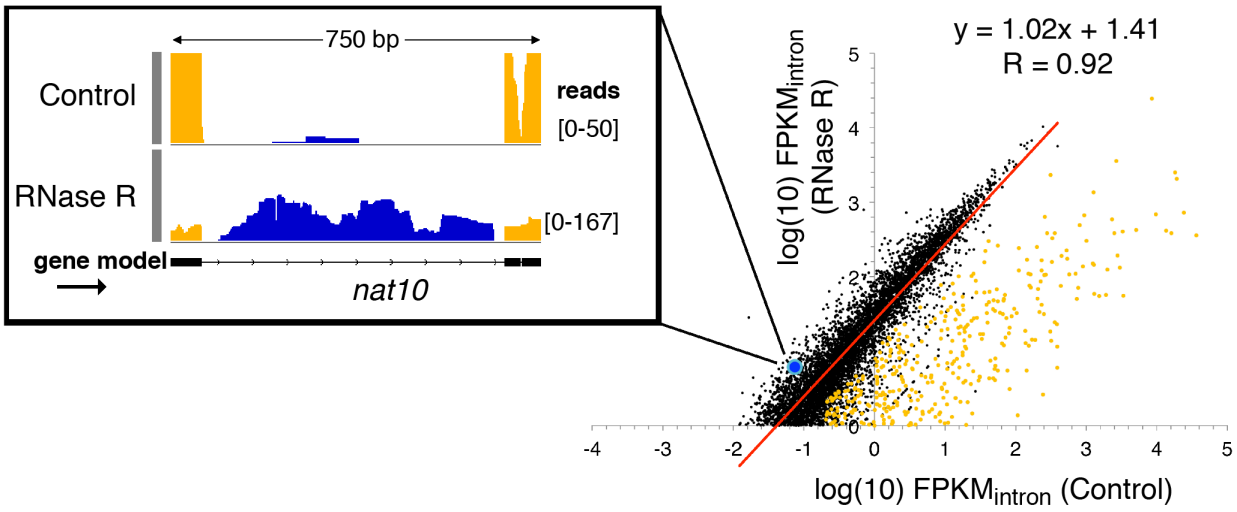
Supplemental Figure S3. Comparison of $\log_{10}$ $FPKM_{intron}$ values for 4417 introns from two independent samples of cytoplasmic RNA. Reproducibility between experiments was high (R = 0.86). These introns were derived from a total of 2963 genes, approximately 30% of all genes with an $FPKM_{exon} \geq 2$.



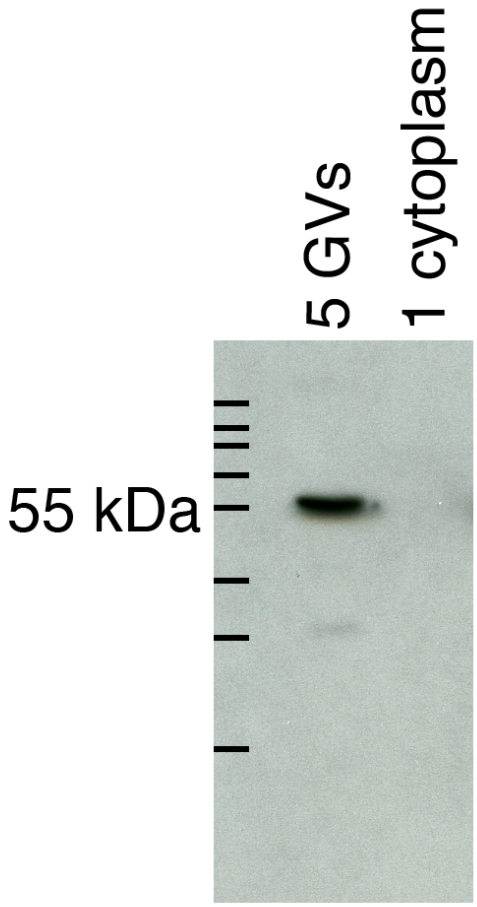
Supplemental Figure S4. All cytoplasmic intronic reads are from the same strand as corresponding exonic reads. Reads were mapped to 14,875 introns in this sample of cytoplasmic RNA. One percent of these (green dots) appeared to have some anti-sense reads. When examined individually, however, each case involved overlapping genes with opposite orientations. The inset shows one such example.



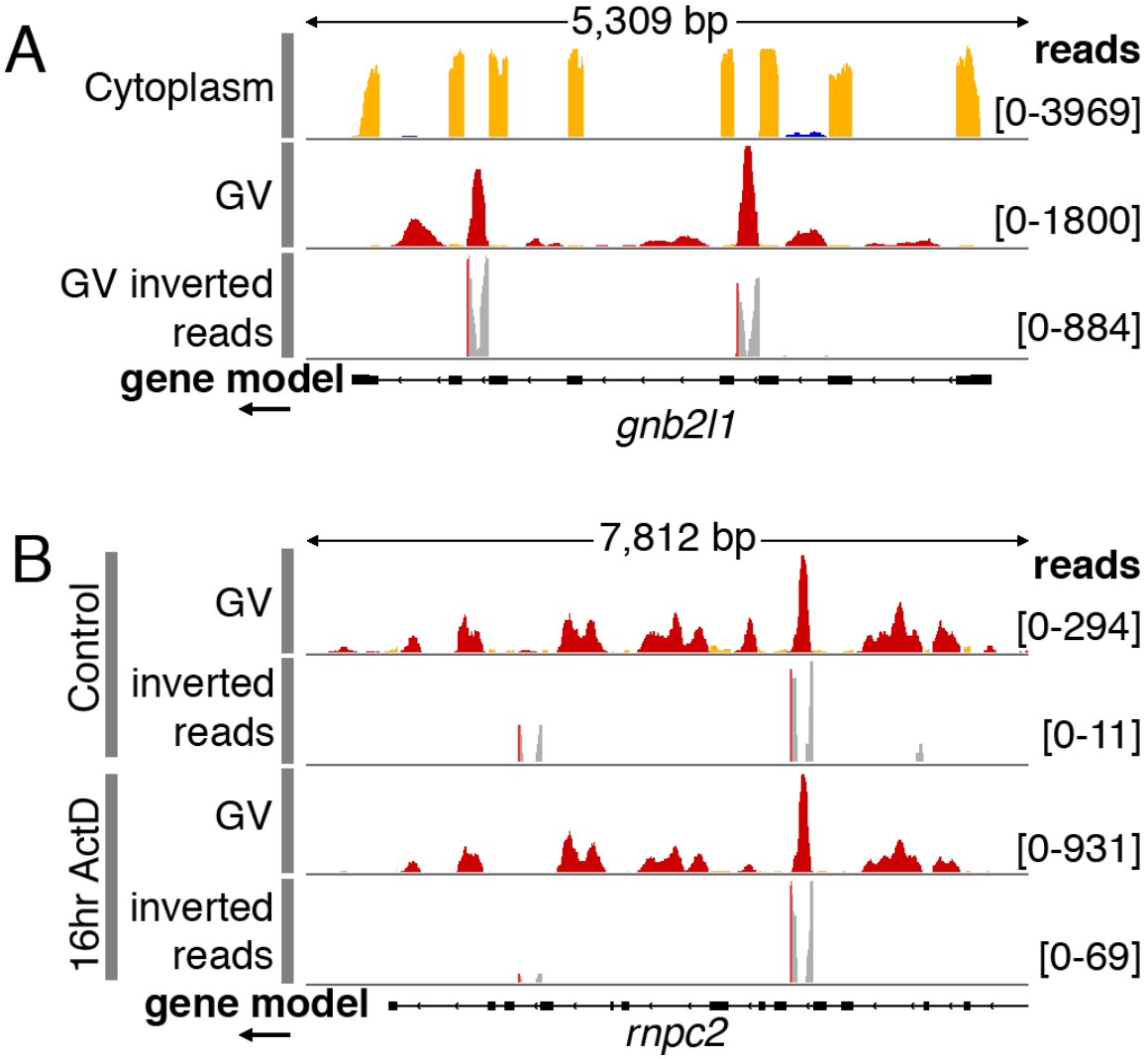
Supplemental Figure S5. Cytoplasmic sisRNA sequences are as stable as mRNA sequences after inhibition of transcription with actinomycin D. Shown is a comparison of $\log_{10}$ $FPKM_{intron}$ / $\log_{10}$ $FPKM_{exon}$ values for cytoplasmic RNA from control and actinomycin-treated oocytes. Correlation between the two samples is high (R = 0.83). Each point represents the ratio between the $\log_{10}$ FPKM value for a specific intron and the $\log_{10}$ FPKM value for the entire mRNA. The inset shows an example: the intron value is based on the blue fragments and the intron length, the exon value on the yellow fragments and the total exonic length.



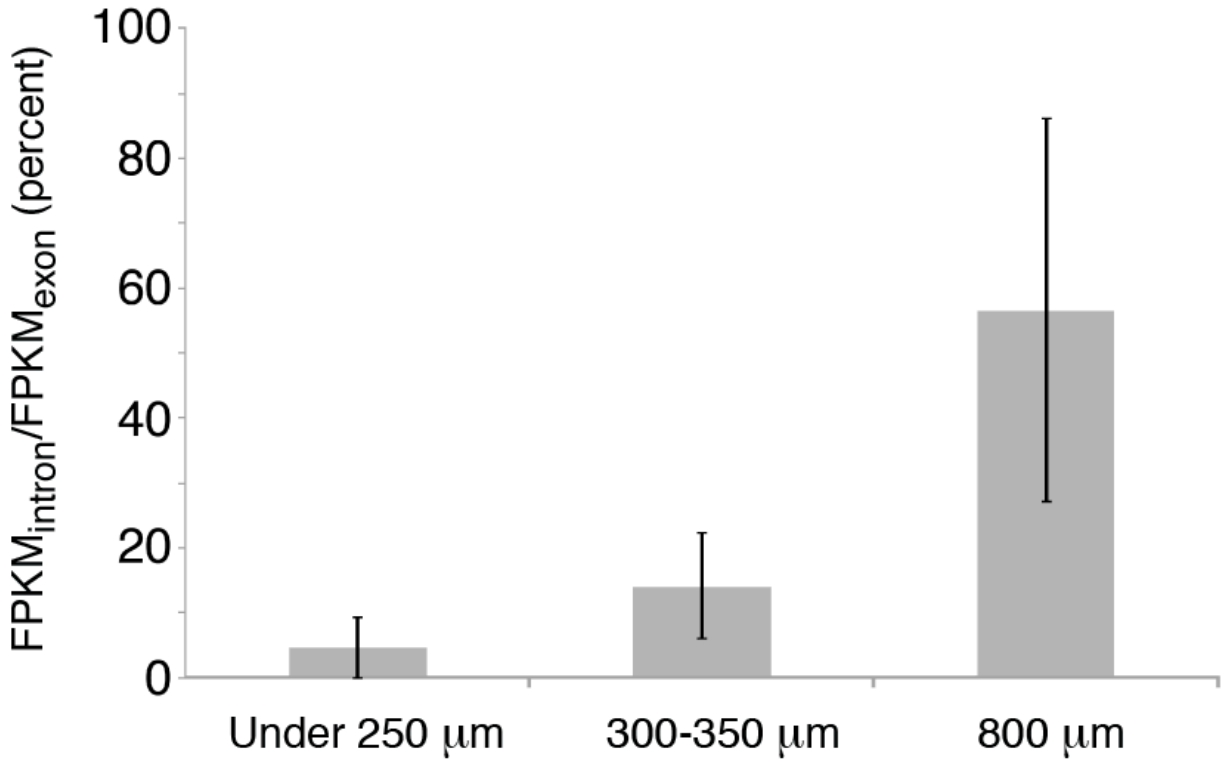
Supplemental Figure S6. Treatment of a cytoplasmic RNA sample with RNase R increases the number of sisRNA reads by 10-100 fold. Log_{10} $FPKM_{intron}$ values for control and RNase R treated samples are plotted. The y-intercept of the regression line is 1.4, equal to an average 25 fold increase in reads after RNase R. Points that fell away from the main cluster (yellow) were examined individually and in every case corresponded to alternative splicing, not genuine sisRNA sequences. Inset shows the increase in reads (blue) for a typical sisRNA (in the *nat10* gene).



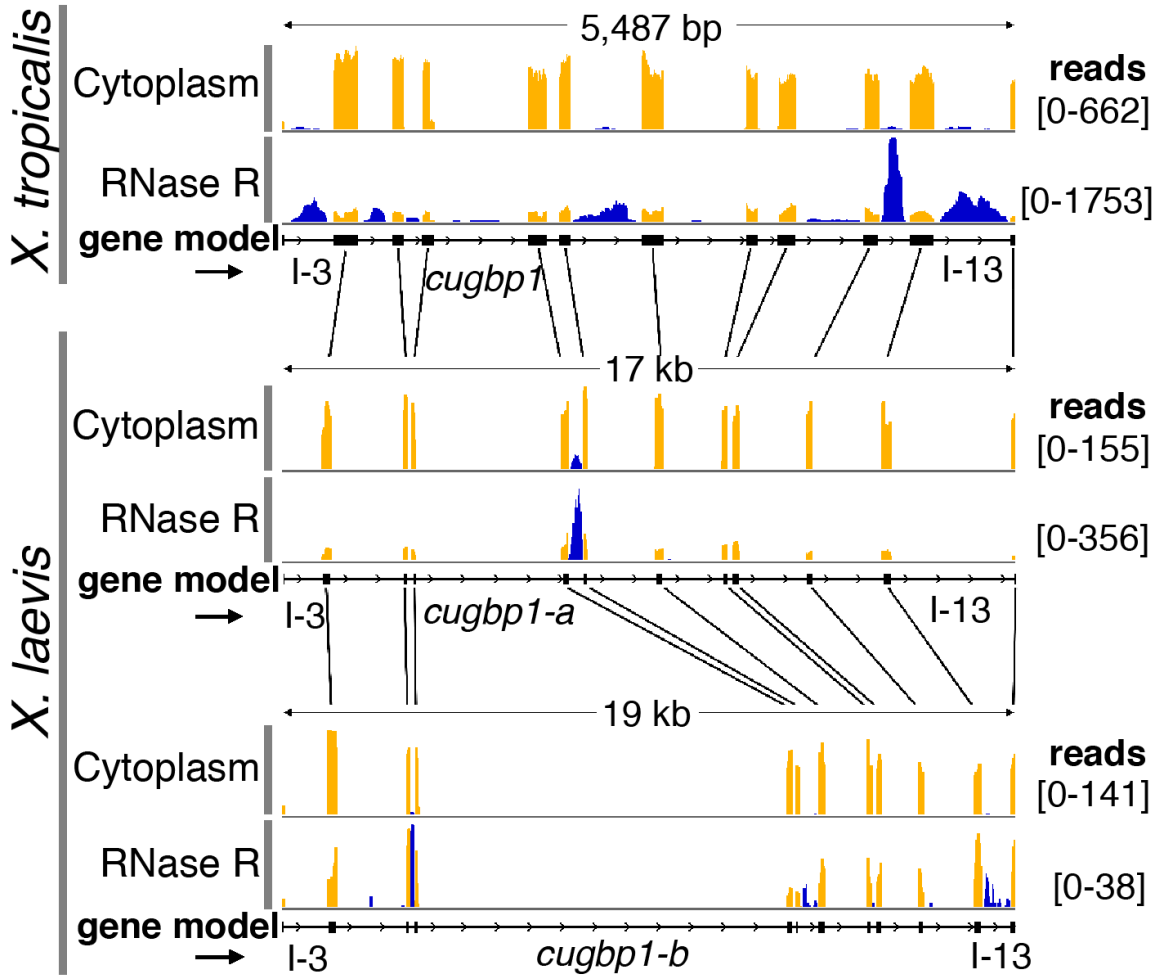
Supplemental Figure S7. The RNA lariat debranching enzyme Dbr1 is present in the GV but is not detectable in the cytoplasm of the *X. tropicalis* oocyte. Shown here is a western blot of proteins from 5 GVs or 1 cytoplasm probed with an antibody against human Dbr1.



Supplemental Figure S8. Some nuclear sisRNAs are stable lariats. *(A)* Inverted reads with mismatches at the junction can be found for some nuclear sisRNA molecules by searching the pool of unmapped reads (Taggart et al. 2012). These sisRNA molecules are presumably lariats. The sisRNAs that lack inverted reads could still be lariats, since inverted reads are produced only if the reverse transcriptase can pass through the branchpoint. *(B)* All nuclear sisRNAs are equally stable after transcription has been inhibited by actinomycin D. The persistence of lariat sisRNAs after actinomycin treatment shows that they are not transient splicing intermediates.



Supplemental Figure S9. The abundance of cytoplasmic sisRNAs increases relative to mRNA during oogenesis. The graph shows the averages for the 15 most abundant cytoplasmic sisRNAs.



Supplemental Figure S10. Cytoplasmic sisRNAs do not occupy the same introns in orthologous genes of *X. tropicalis* and *X. laevis*. The top two tracks show that the *cugbp1* gene (introns 3-13) of *X. tropicalis* gives rise to multiple cytoplasmic sisRNAs (blue), with the most prominent in intron 12. The two homologous genes in *X. laevis* (*cugbp1-a* and *cugbp1-b*) differ from each other and from the *X. tropicalis* gene. The most prominent sisRNA comes from intron 7 in *cugbp1-a* (middle tracks) but from intron 5 in *cugbp1-b* (bottom tracks).

EXPERIMENTS		Orientation	arfgap2(1)	arfgap2(2)	eif4a1	faf2
detection	mRNA	forward	AACCCGCATCTACCACAGTTACCA	CGCCAAAGCCATTTCCTCAGACAT	AGATTACATGGGTGCCTCTTGCCA	TCGAATCAATGGACCAATGTCGCC
		reverse	AAGTTCCTGATACGCCAGCCTCAT	TCACTCCATTAGCCAGAACTGCCA	TCAACATCTCGTCGGCTTCATCCA	GCTGTTAAGGAAATGGGTGACCTC
	sisRNA	forward	TACGCGGGCAGGGACTAATAACAA	CAGCCAGCCAGCCAATTATTTCCA	TTGGAATGAGGCTGTTTATGGGCG	ATTGATTAACCAGCAGCCTGAAGA
		reverse	AGAGTGCACACTTGCTCTCCTTGA	TTCATCAAGAGGTCTAGGGCAGCA	ACTCTGGCCCACAGAGCTTACAAT	TATGCTCAGACCTGCTGTTGCACT
in vitro transcription	pre-mRNA	forward	AACCCGCATCTACCACAGTTACCA	CGCCAAAGCCATTTCCTCAGACAT	AGATTACATGGGTGCCTCTTGCCA	TATTAACCCTCACTAAAGGACTTGACAGGTATCGAATCAA
		reverse	AAGTTCCTGATACGCCAGCCTCAT	TCACTCCATTAGCCAGAACTGCCA	TCAACATCTCGTCGGCTTCATCCA	TTTTTTTTTTTTTTTTTTTTCCTTGGTTGTGGCCGAGACACCACATA
construct	sisRNA	forward	ATTAACCCTCACTAAAGGAACTGGAGCTGGTTCCAACTGAGA	ATTAACCCTCACTAAAGGAGACGCCAGTTGCTAACAAGTCCT	ATTAACCCTCACTAAAGGATCCAGCAGCGAGCTATTATGCCT	ATTAACCCTCACTAAAGGGGACGAACGTTCCTTTGCTTCTGA
		reverse	GGATCCAGGACTTGTTAGCAACTGGCGTCT	GGATCCAGCCGGCCATATACGCAATTTCAC	GGATCCTGAACTTCTTGGTCACCTCGAGCA	GGATCCTTTCTTGGTCAGCACGCAGAGATG
lariat detection	sisRNA	forward	CCTTTCAGTTGAGGAGGCGA	CCCCGCTAACTTGACCCATT	CTTATATCTGGCGGGGGCAG	AGTGCAACAGCAGGTCTGAGCATA
		reverse	GGACTTCAGGACACTCGCAA	GGCAGCAGGGAAAACCACTA	GTACCGCCCATAAACAGCCT	TCTTCAGGCTGCTGGTTAATCAAT
inverted *in vitro* transcription construct	sisRNA -piece1	forward	TAATACGACTCACTATAGGGGCAGCCATGTGATGCCCTATA	TAATACGACTCACTATAGGGAAATGTAACCTGCTGTGGGTCGC	TAATACGACTCACTATAGGGGAGGAATCTGCCATAGAAGTGCCT	ATTAACCCTCACTAAAGGGAACGTACCTTTACCTTTCTCTGTTAA
		reverse	GAAATGGACCCACTCATGAGCAAGTGGTGCCCCAAATG	GCCGTTCTACTGACTCACTTGTGCCCAAACATGCAAATCA	GCTCATGATACTCACTGAGTGCTGCACTGGCGCCACG	GGTTCAGGTACATACTCATCAGAGAACAGTGTACTCTGTA
	sisRNA -piece2	forward	GGGCACCACTTGCTCATGAGTGGGTCCATTTCCATTCTGTATA	GTTTGGGCACAAGTGAGTCAGTAGAACGGCCAACTGG	CCAGTGCAGCACTCAGTGAGTATCATGAGCCTGCCTCTG	AGTACACTGTTCTCTGATGAGTATGTACCTGAACCAGCAGGGAA
		reverse	GTAGGGCTCAGACTGAAGGC	AACGGAAGAGATCGGTGGCCA	TTTGCTGATGGGTAGTGAGCAT	ACACACAGTTAGTCATATTGCAAA

Supplemental Table S1. Primers used for RT-PCR experiments. T3 and T7 promoter sequences in primers are underlined.