

Additional File 1 for:

ALG-1 influences accurate mRNA splicing patterns in the *C. elegans* intestine and body muscle tissues by modulating splicing factor activities.

Kotagama K, Schorr AL, Steber HS, Mangone M*.

* To whom correspondence should be addressed. Tel: (480) 965-7957; Email: mangone@asu.edu

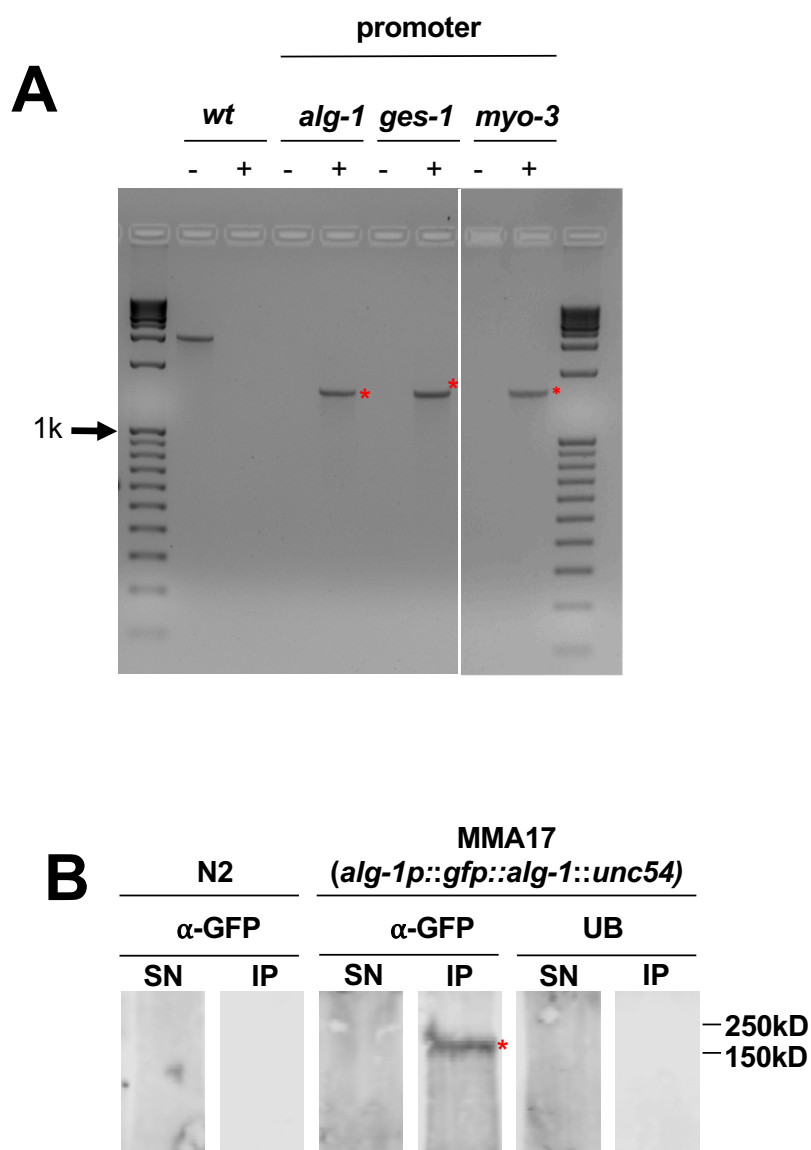
This PDF includes:

Figures S1-S10

Table of Contents

Additional File 1

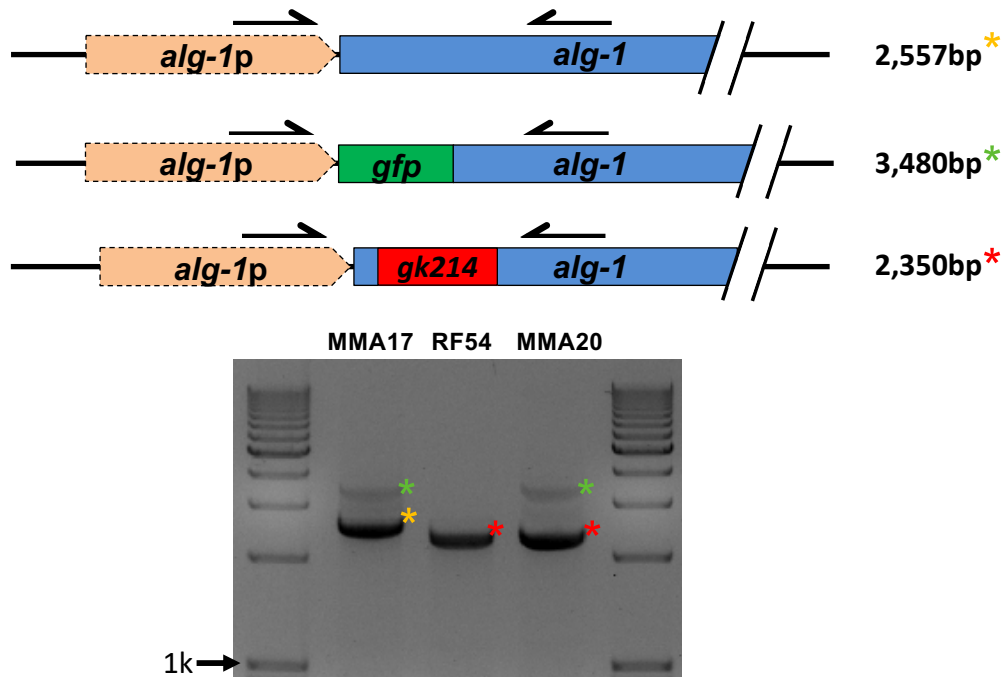
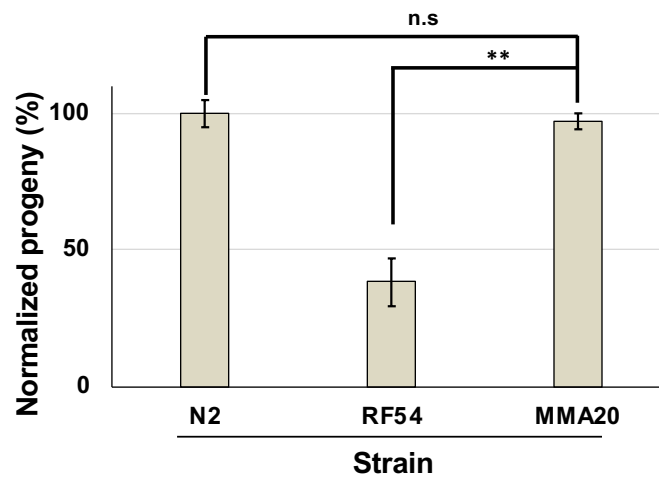
Figure S1: Validation of transgenic strains for miRNA-target identification.....	3
Figure S2: Genotyping and layoff experiment to validate <i>in vivo</i> the functionality of the GFP tagged ALG-1 construct.....	4
Figure S3: Summary of the sequencing results from our tissue specific ALG-1 immunoprecipitation.....	5
Figure S4: Comparison of our tissue specific ALG-1 pull-down to previously identified <i>alg-1</i> RNA (<i>Zisoulis et al.</i> , 2010).....	6
Figure S5: Validation of selected hits detected in our tissue specific ALG-1 pull-down	7
Figure S6: Comparative analysis of our tissue specific ALG-1 pull-down to previously identified intestine and body muscle miRNA (mime-seq) (<i>Alberti et al.</i> , 2018).....	8
Figure S7: Layoff experiment to determine the efficacy of RNAi targeting splicing factors.....	9
Figure S8: <i>lin-10</i> , <i>unc-52</i> and <i>ret-1</i> exon skipping events in the intestine (1).....	10
Figure S9: <i>lin-10</i> , <i>unc-52</i> and <i>ret-1</i> exon skipping events in the intestine (2).....	11
Figure S10: Gene Ontology analysis of genes with 2-fold splicing change in ALG-1 or ALG-2 knockout worms (Supplementary Table S4).....	12



Supplemental Figure S1: Validation of transgenic strains for miRNA-target identification. A) Agarose gel electrophoresis of PCR products obtained from genomic DNA using single-worm PCR approach from transgenic *C. elegans* strains prepared in this study. Each strain was tested using primers that detect either the presence (+) or the absence (-) of the integrated MosSCI transgene. Red asterisks mark expected bands confirming integration of our constructs. B) Western blot experiments from *wt* N2 worms or endogenous ALG-1 pull-down construct performed using either α -GFP antibodies or unconjugated beads (UB) on immunoprecipitated protein obtained from each strain. (IP) immunoprecipitated protein, (SN) supernatant. Red asterisk marks the expected band.

A

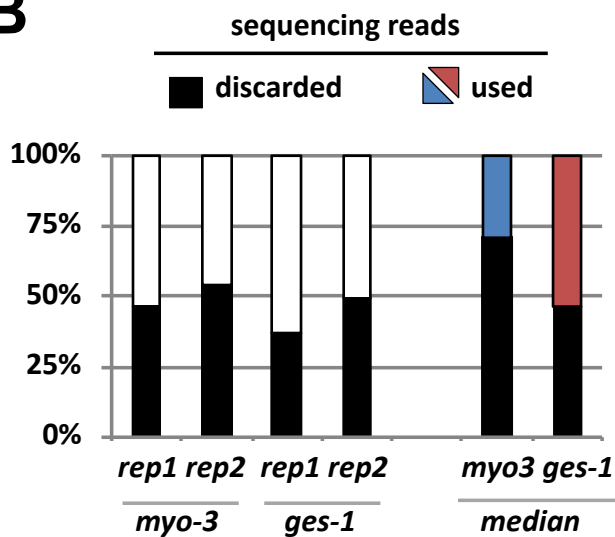
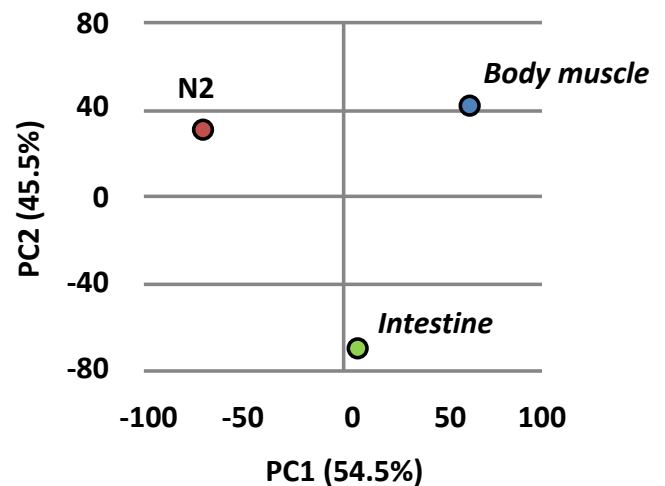
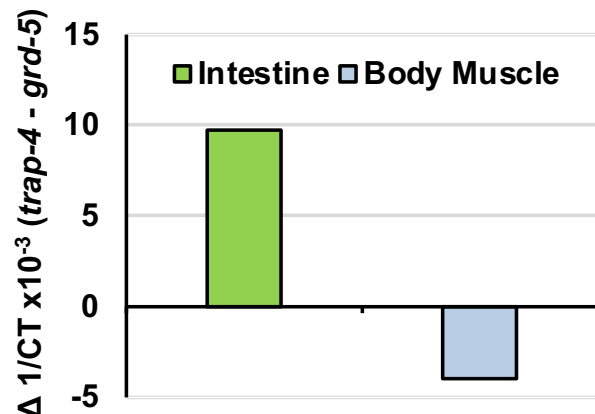
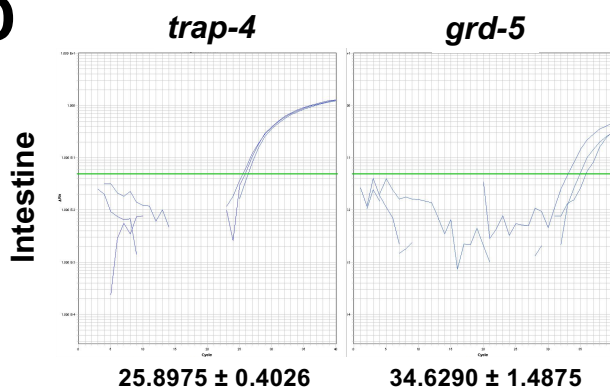
Genotypes:

MMA17 - [*alg-1p:gfp:alg-1:unc-54*] IIRF54 - *alg-1(gk214)* XMMA20 - *alg-1(gk214)* X; [*alg-1p:gfp:alg-1:unc-54*] II**B**

Supplemental Figure S2: Genotyping and layoff experiment to validate *in vivo* the functionality of the GFP tagged ALG-1 construct. A) Agarose gel visualizing single worm genomic PCRs testing the presence of the *gk214* mutation before and after the cross. Yellow asterisks indicate the genomic locus with *wt alg-1* (2,557bp). Green asterisks indicate *gfp* tagged transgenic *alg-1*. Red asterisks indicate the presence of the *gk214* mutation (2,350bp). In the MMA20 strain we re-introduced our cloned *alg-1* to the RF54 strain, which lacks of endogenous *alg-1*. B) Results from the layoff experiment testing fertility. The loss of fertility seen in the *alg-1* knockout strain RF54 is rescued in the MMA20 strain which expresses our GFP tagged transgenic ALG-1, suggesting that our ALG-1 pull-down construct is functional and mimic endogenous *alg-1* activity. One-way ANOVA. ** $p < 0.01$. $n = 5$. Error bars indicate the standard error of the mean.

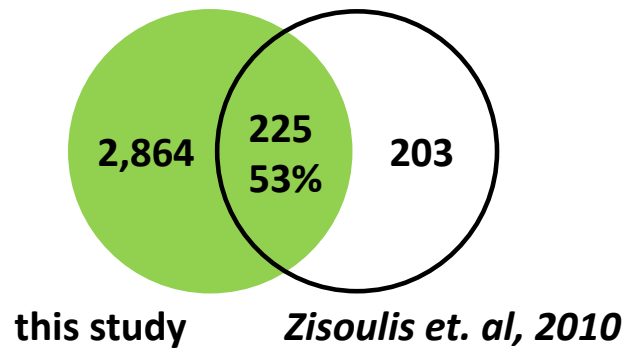
A

Sample		Sequencing reads		
		Total (#)	Mapped (#)	Mapped (%)
<i>alg-1p</i>		23,277,135	18,292,992	78.6
Intestine (<i>ges-1p</i>)	Experiment (<i>rep 1</i>)	31,424,437	23,241,554	73.9
	Replicate 2 (<i>rep 2</i>)	23,522,056	17,861,605	75.9
Body muscle (<i>myo-3p</i>)	Experiment (<i>rep 1</i>)	24,092,664	21,327,095	88.5
	Replicate 2 (<i>rep 2</i>)	25,347,564	22,062,749	87.0
N2		21,894,026	19,378,018	88.5

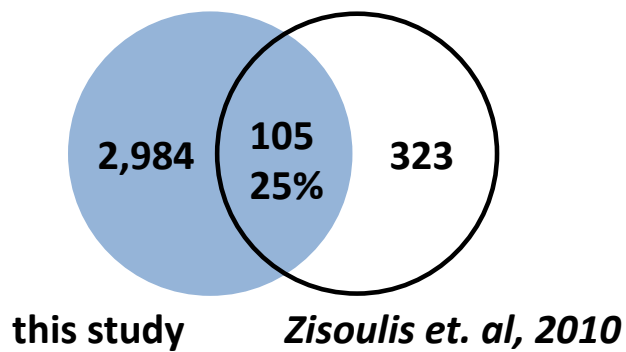
B**C****D**

Supplemental Figure S3: Summary of the sequencing results from our tissue specific ALG-1 immunoprecipitations. A) Overview of the sequencing reads obtained for each immunoprecipitation. Each sample was sequenced in duplicate with biological replicates. We obtained at least 15M reads each dataset mapped to the *C. elegans* genome (WS250). B) In this study we have used only the top ~50-75% positive hits produced by the Cufflinks algorithm. . Blue; Body muscle. Red; Intestine C) Principal Component Analysis (PCA) between the median *fpkm* values among replicates shows differences between tissues within our datasets. Each dot represents a median experiment and replicate. D) qPCR analysis validation of sequencing results from tissue-specific ALG-1 immunoprecipitations. qPCR experiments were performed in triplicate using cDNA generated from mRNA extracted by ALG-1 immunoprecipitation in the intestine and body muscle tissues. We tested *trap-4* (abundant in the intestine dataset and not detected in the body muscle tissue) and *grd-5* (abundant in the body muscle dataset, and not in the intestine tissue) (Supplemental Table 1). The line graphs represent amplification plots of the qPCR data of the intestine IPs. The numbers below the graph show the cycle threshold (CT) value and its standard deviation. The bar graph on the right illustrates the relative change in abundance of *trap-4* and *grd-5* between the intestine and body muscle IPs and is in agreement with our sequencing results.

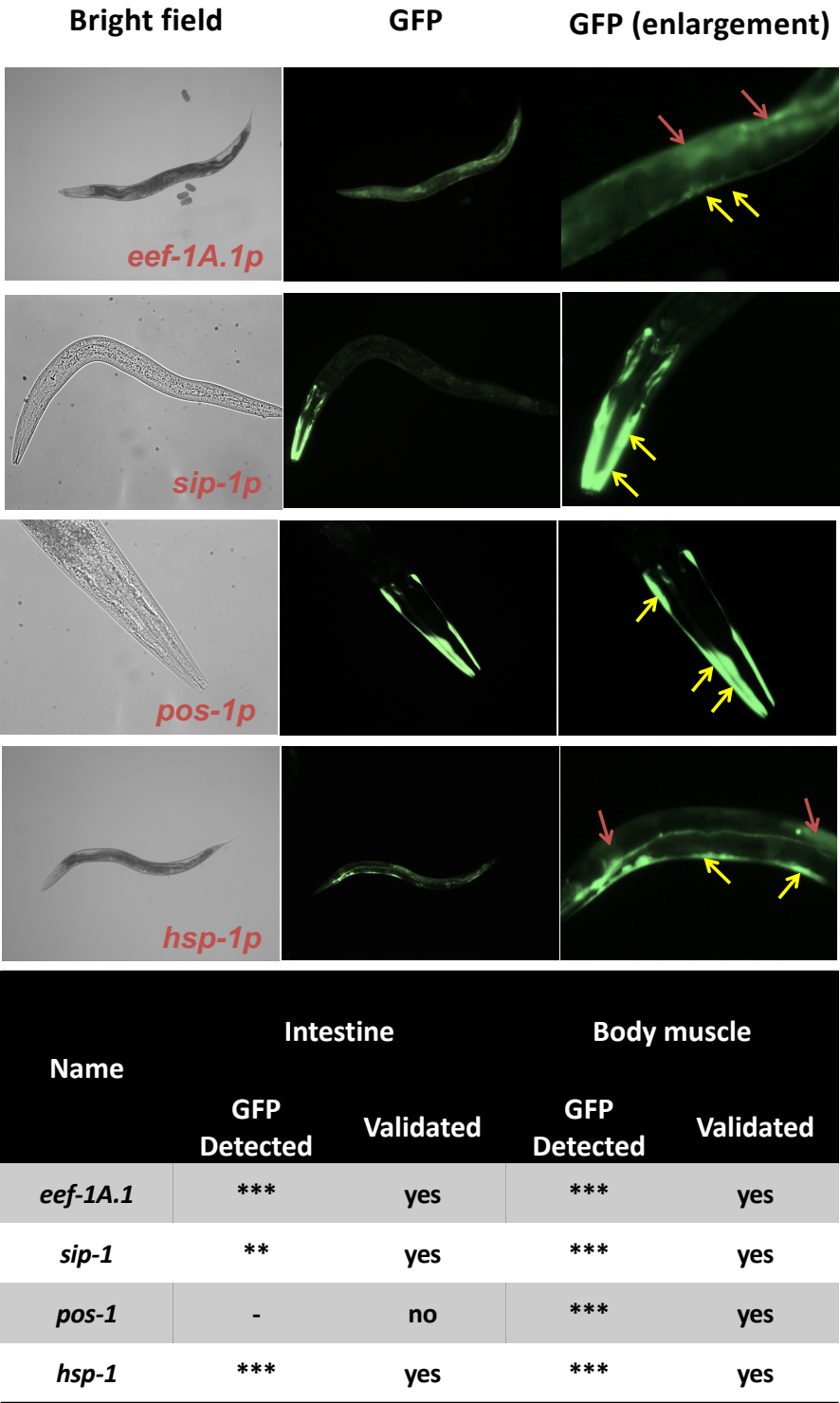
Intestine



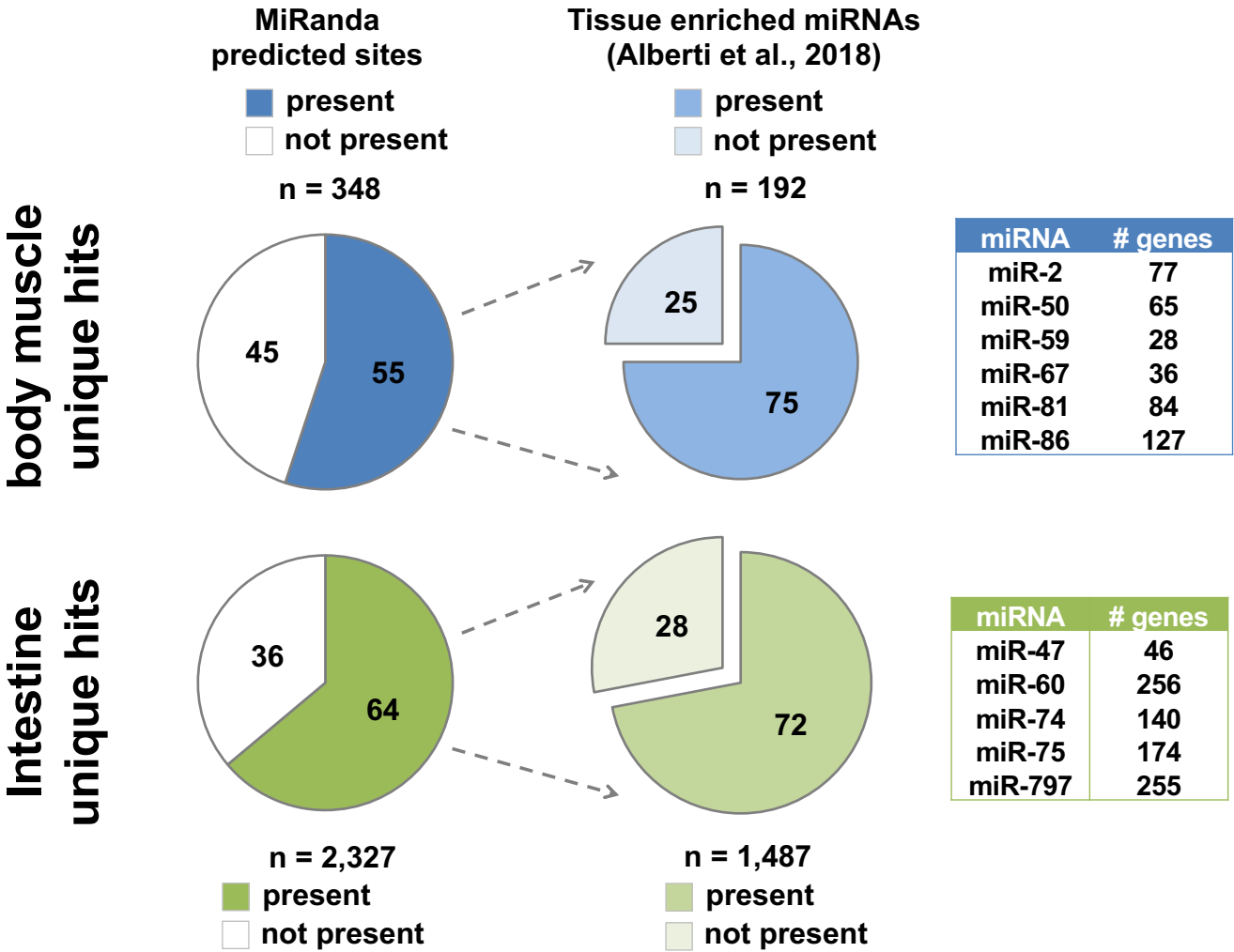
Body muscle



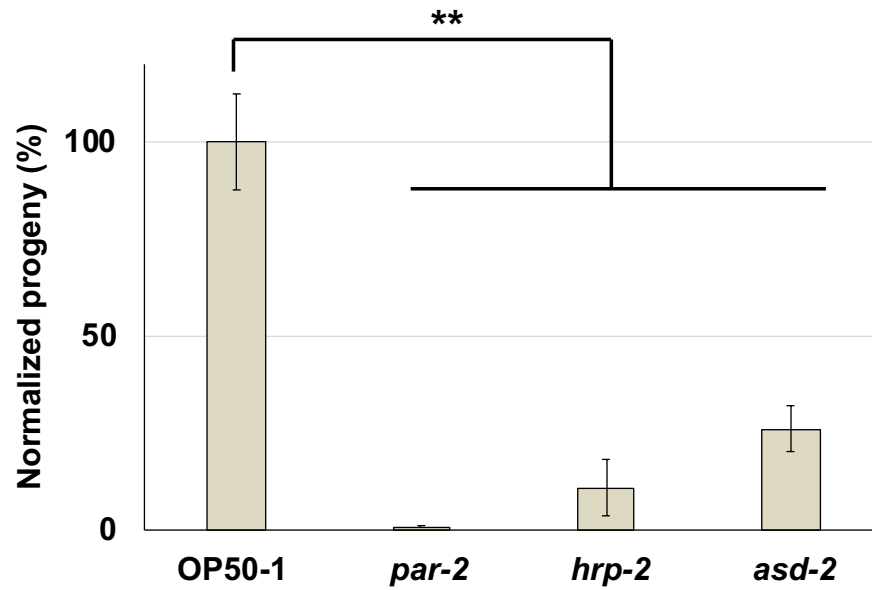
Supplemental Figure S4: Comparison of our tissue specific ALG-1 pull-down to previously identified *alg-1* RNA (Zisoulis et al., 2010). Venn diagram showing the comparison of hits between our datasets and a comprehensive RNA pull down of miRNA targets.



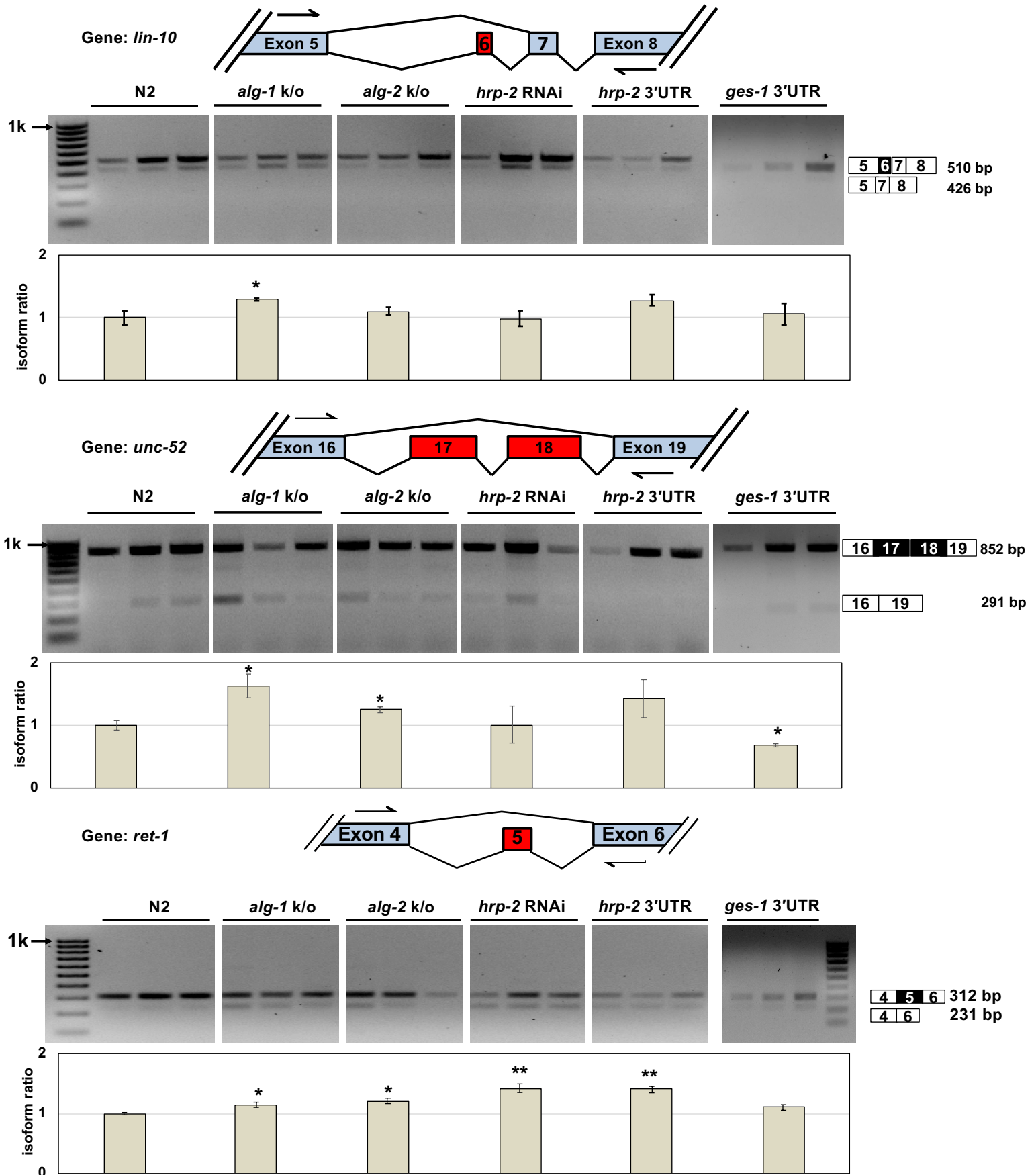
Supplemental Figure S5: Validation of tissue localization selected hits detected in our tissue specific ALG-1 pull-down. We have cloned selected promoters from genes identified in our tissue specific ALG-1 pull-down, and fused it to GFP reporters to validate the expression localization pattern. Except for one case, we detected strong GFP expression in the expected tissues. Red arrow mark intestine expression, Yellow arrow mark body muscle expression. Putative expression index: - not detected; * low expression; ** expressed; *** strong expression



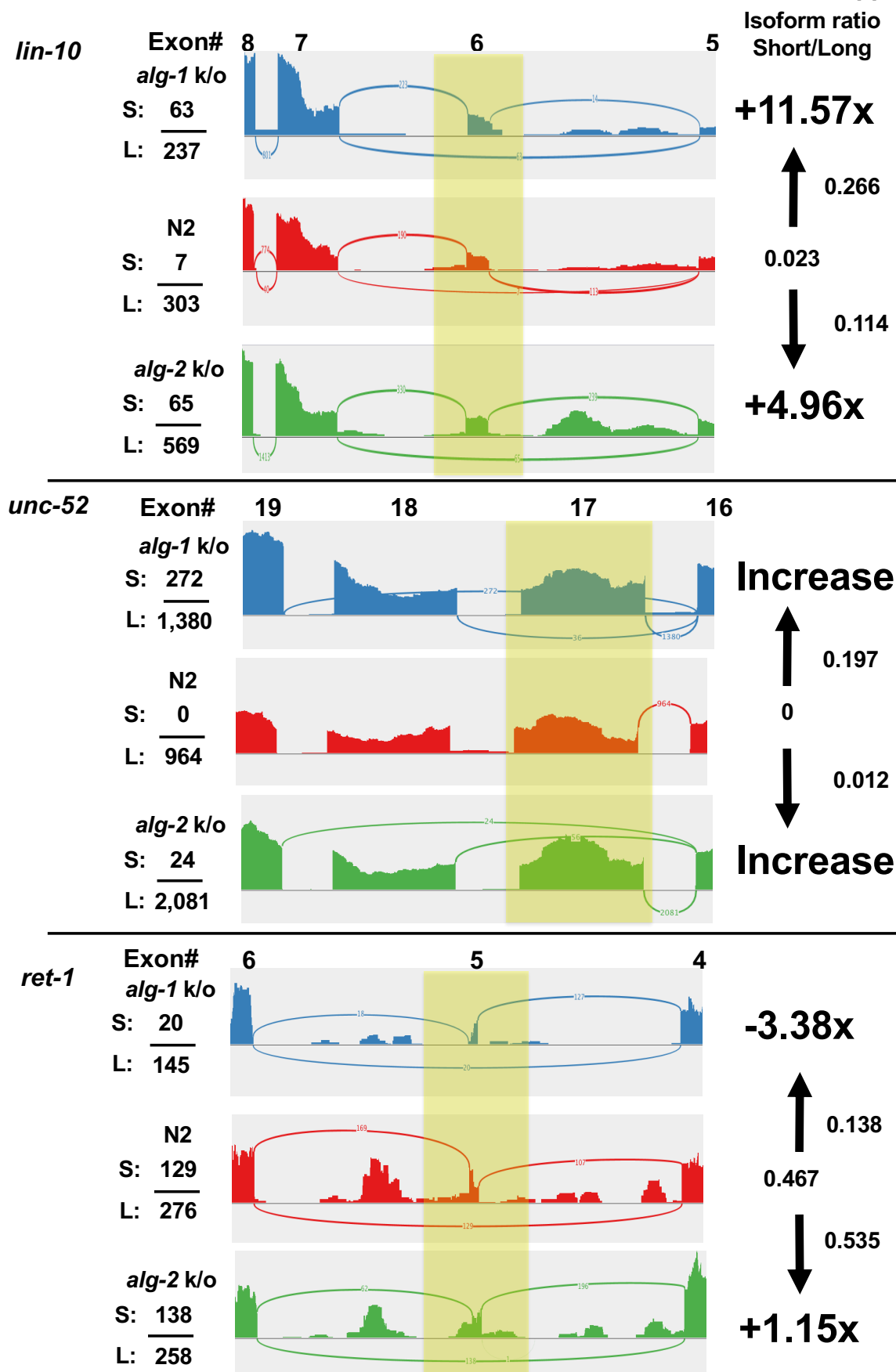
Supplemental Figure S6: Comparative analysis of our tissue specific ALG-1 pull-down to previously identified intestine and body muscle miRNA (mime-seq) (Alberti et al., 2018). The left pie chart shows the percentage of genes identified in our study with predicted miRNA binding sites as for MiRanda algorithm. The right pie chart shows how many of those genes contain a predicted site for intestine or body muscle miRNAs identified in Alberti et al., 2018. The table shows how many genes identified in our study are targeted by unique tissue-specific miRNAs identified in Alberti et al., 2018.



Supplemental Figure S7: Layoff experiment to determine the efficiency of RNAi targeting *hrp-2* and *asd-2*. RNAi performed against the gene *par-2* (positive control) produced almost no hatched larvae as expected. A decrease in the number of progeny was observed when RNAi was performed for the splicing factors *hrp-2* and *asd-2* indicating the knockdown of these genes. ** $p < 0.01$ $n = 5$



Supplemental Figure S8: *lin-10*, *unc-52* and *ret-1* exon skipping events in the intestine (1). Top panels: The *lin-10*, *unc-52* and *ret-1* gene models are shown to illustrate the alternatively spliced exons (red) and the primer pairs used in the RT-PCR experiments to detect intestine specific alternative splicing modulated by miRNAs. Bottom Panels: Agarose gels visualizing RT-PCRs performed in triplicate on independent RNA extractions to test changes of alternative splicing of individual genes. * $p < 0.05$; ** $p < 0.01$



Supplemental Figure S9: *lin-10*, *unc-52* and *ret-1* exon skipping events in the intestine (2) Sashimi plots generated by re-analyzing RNA-seq data from Brown et. al 2017. Replicates for each strain were combined and analyzed with the TopHat bioinformatic package. The plots illustrate the number of reads mapped for each splice junction. Top and middle panels: The splicing events for *lin-10* and *unc-52* showed an increase of exon skipping occurring in both miRNA deficient strains. Lower panel: *ret-1* exon 5 skipping increases slightly in *alg-2* but shows increased exon inclusion in *alg-1* knockout worms. Yellow boxes highlight alternatively spliced exons analyzed in this figure S: total reads for the short isoform; L: total reads for the long isoform.

Biological Process	Fold Enrichment	p-value
mesoderm development	2.12	2.73E-03
positive regulation of multicellular organism growth	2.01	1.42E-03
male sex differentiation	1.99	1.25E-05
nematode male tail tip morphogenesis	1.97	9.23E-05
muscle structure development	1.91	1.44E-04
body morphogenesis	1.9	5.05E-04
tissue development	1.88	6.63E-07
regulation of supramolecular fiber organization	1.85	2.70E-03
regulation of MAPK cascade	1.85	1.19E-03

Supplemental Figure S10: Gene Ontology analysis of genes with 2-fold splicing change in ALG-1 or ALG-2 knockout worms (Supplementary Table S4). The results suggest that genes involved with development and tissue morphogenesis are miss-spliced in worms deficient in the miRNA pathway.