Regulated Alternative Splicing of *Caenorhabditis elegans rsp* mRNAs

By

Benjamin Schmidt

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2012

Date of final oral examination: 6/8/12

The dissertation is approved by the following members of the Final Oral Committee:

Philip Anderson, Professor of Genetics

Michael Culbertson, Professor of Genetics and Molecular Biology

Audrey Gasch, Associate Professor of Genetics

Scott Kennedy, Associate Professor of Genetics and Pharmacology

Allen Laughon, Professor of Genetics and Medical Genetics

This thesis is dedicated to my parents, Paul and Debbie.

Acknowledgments

The budding scientist receives a slew of support along the way:

- Phil provided the initial enthusiasm for a project that got me excited about splicing. Allowing me to ask the questions that most fascinated me ensured I never lost interest. I leave with more questions regarding how and why than when I started. To me, that's the mark of a job well done.
- The rest of my committee, Mike Culbertson, Audrey Gasch, Scott Kennedy, and Al Laughon, have provided helpful guidance and suggestions throughout the years. Our discussions were always pleasant, and I found myself leaving meetings with renewed enthusiasm to pursue new experiments and questions.
- Anderson lab members have shared the good, the bad, and the ice cream. Jasmine, Lisa, and Leah were with me at the start, Amy joining not long after. They answered the naïve questions of a first-year, suggesting how to do experiments and teaching me the art of *C. elegans* genetics. One of their most important tips was that the lab could not function without Bonnie, who keeps us stocked with media, helpful suggestions, and a pleasant lab environment. Virginia has provided renewed interest in what NMD is doing and why. I wish her the best of luck in the years to come.
- The Kennedy lab has provided excellent suggestions and critiques in our many lab meetings. Their willingness to share media and cycler time has been invaluable.
- My family has provided a constant source of encouragement and love. I would not be typing this were it not for you all.

Table of Contents

Abstract

Alternative splicing is common amongst eukaryotes and is regulated in part by a large number of proteins. Two large protein families have emerged as general regulators of alternative splicing in eukaryotes: serine/arginine-rich (SR) proteins and heterogeneous ribonucleoproteins (hnRNPs). mRNAs that encode SR and hnRNP proteins are often alternatively spliced and this thesis focuses on the *regulated* alternative splicing of mRNAs that encode SR proteins in the nematode *Caenorhabditis elegans*.

I found that six of the seven *C. elegans rsp* mRNAs (known as SR mRNAs in other eukaryotes) are alternatively spliced. Such splicing leads to one of two types of mature *rsp* mRNAs: (i) *rsp* mRNAs that encode functional RSP proteins, termed *rsp(+)* mRNAs, or (ii) *rsp* mRNAs that introduce premature termination codons (PTC), termed *rsp(PTC)* mRNAs. PTC-containing mRNAs are substrates for the nonsense-mediated mRNA pathway (NMD) and are rapidly degraded in wild type.

I investigated if and how *rsp* splicing patterns are regulated in three ways. (i) I investigated if *rsp* splicing patterns are developmentally regulated. Early in development, a drastic decrease in several *rsp* mRNAs occurs that correlates with an increase in the proportion of *rsp(PTC)* mRNAs. I suggest the increase in the proportion of *rsp(PTC)* mRNA is a means to remove excess *rsp* mRNA since PTC-containing mRNAs are rapidly degraded in wild type. (ii) I investigated if RSP proteins autoregulate their own splicing patterns. I found overexpression of an RSP protein affects the splicing pattern of its own mRNA by increasing the proportion of *rsp(PTC)*, and reduction in RSP protein levels decreases the proportion of *rsp(PTC)* mRNA. (iii) I investigated if one or more RSP proteins can regulate

the splicing pattern of a different *rsp* pre-mRNA. I found RSP-2 and RSP-4 regulate the splicing pattern of *rsp-5* mRNA, and RSP-4 regulates the splicing pattern of *rsp-7* mRNA.

I hypothesize these regulated splicing events are examples of post-transcriptional gene regulation. Splicing to increase or decrease the proportion of *rsp(PTC)* mRNA is therefore a means to fine-tune the levels of *rsp(+)* mRNAs.

Abbreviations

SR or RSP: serine/arginine rich protein

hnRNP: heterogeneous ribonucleoprotein

hnRNA: heterogeneous ribonucleic acid

PTB (also known as hnRNP I): polypyrimidine-tract binding protein

snRNP: small nuclear ribonucleoprotein

snRNA: small nuclear ribonucleic acid

ESE: exonic splicing enhancer

ISE: intronic splicing enhancer

SELEX: systematic evolution of ligands by exponential enrichment

RRM: RNA recognition motif

RS domain: arginine/serine rich domain

KH: K-homology domain

Pol II: RNA Polymerase II

CTD: COOH-terminal Domain of Pol II

H3K36me3: histone 3 lysine 36 trimethylation

H3K9me2: histone 3 lysine 9 dimethylation

H3K27me3: histone 3 lysine 27 trimethylation

CTCF: CCCTC-binding factor

PTC: premature termination codon

NMD: nonsense-mediated mRNA decay

UPF: up-frameshift

smg: suppressor with morphogenic effects on genitalia

EJC: exon junction complex

PABP: poly(A)-binding protein

ORF: open reading frame

uORF: upstream open reading frame

RUST: Regulated unproductive splicing and translation

RNAi: RNA interference

siRNA: short interfering RNA

GFP: green fluorescent protein

bp: base pair

nt: nucleotide

PCR: polymerase chain reaction

RT-PCR: reverse transcription polymerase chain reaction

qRT-PCR: quantitative reverse transcription polymerase chain reaction

Chapter 1: An Introduction To Alternative Splicing And Nonsense-Mediated mRNA Decay (NMD)

Eukaryotic mRNAs undergo a series of processing events including transcription, splicing, capping, polyadenylation, export, translation, and decay. Each of these processes is tightly regulated and can have profound impacts on gene expression. My thesis focuses on two of these processes: patterns of alternative splicing and nonsense-mediated mRNA decay (NMD).

Alternative splicing generates multiple mature mRNAs from a common pool of pre-mRNA by altering exons included in mature transcripts. Alternative splicing expands the coding potential of the pre-mRNA pool, as translation of different combinations of protein-encoding exons yields different protein isoforms. Alternative splicing can also generate mature mRNAs containing premature termination codons (PTCs), which subsequently are substrates for NMD and degraded.

Identification of *Caenorhabditis elegans* alternatively spliced *rsp* mRNAs that contain PTCs led me to hypothesize that alternative splicing of *rsp* pre-mRNAs is regulated. I suspected that regulation of *rsp* alternative splicing might regulate *rsp* gene expression. The role of RSP proteins in regulating alternative splicing of *rsp* mRNAs is the central theme of this thesis. I review and discuss in this chapter processes of alternative splicing and the factors that regulate splicing. I also give a brief overview of the mechanism of NMD and describe types of mRNA substrates degraded by NMD. Finally, I will discuss connections between alternative splicing and NMD.

Alternative Splicing

In most eukaryotes, removal of introns from pre-mRNAs is a necessary step to generate mature mRNAs. Most genes of higher eukaryotes contain at least one intron, but many undergo extensive splicing (reviewed in [1]). Patterns of alternative splicing are commonly observed in pre-mRNAs that contain more than one intron.

Alternative splicing is a common and necessary mechanism to generate protein diversity in higher eukaryotes. This is true both in mammalian species and in commonly used model organisms. High-throughput sequencing of eukaryotic transcriptomes demonstrates that approximately 95% of human genes with more than one intron are alternatively spliced [2-4], 60% of such genes in *Drosophila* [5,6], 42% in *Arabidopsis* [7], and 25% in *C. elegans* [5,8].

Exons are classified by the ways in which they are incorporated into mature mRNAs (Figure 1.1A). 'Constitutive exons' are always included in mature mRNAs and are more common than alternative exons. Splicing at alternative 5’ and 3’ splice sites can cause constitutive exons to be longer or shorter. 'Cassette exons' are alternative exons contained within alternatively spliced introns, and their inclusion is often regulated. Many pre-mRNAs contain multiple cassette exons, and if such exons never appear in the same mature mRNA, they are termed 'mutually exclusive' cassette exons. Some alternatively spliced introns are completely retained in the mature mRNA. Higher eukaryotes are rife with examples of each of these types of alternative splicing, but the precise mechanism of alternative splicing is not well understood.

Regulation of Alternative Splicing

How alternative splicing is regulated is a long-standing question. A substantial number of proteins play specific roles in regulating alternative splicing, but two large families of proteins, the serine/arginine-rich (SR) proteins and the heterogeneous nuclear ribonucleoproteins (hnRNPs), have emerged as general regulators of alternative splicing (reviewed in [1]). SR proteins generally are positive regulators of alternative splicing and promote inclusion of specific additional exons into mature mRNAs. hnRNPs generally act in an opposite manner and promote exclusion of specific exons in mature messages (Figure 1.1B). However, examples of SR and hnRNP proteins both promoting and inhibiting exon inclusion have been observed. SR proteins are able to antagonize hnRNP interactions by competition for binding to pre-mRNAs.

SR and hnRNP proteins can affect splicing patterns in a concentration dependent manner [9]. Higher concentrations of SR proteins in general correlate with an increase in the use of downstream 5' splice sites (Figure 1.2, and described further in the next section). Higher concentrations of hnRNP proteins generally affect 5' splice site selection in an opposite manner. Thus, stoichiometry of splicing regulators may play important roles in regulation of alternative splicing patterns.

Recent studies suggest that alternative splicing is more complicated than simple regulation by splicing factors. Most splicing occurs co-transcriptionally [10], and rates of transcription influence whether alternative exons are included or excluded from mature mRNAs [11]. Moreover, nucleosome positions correlate with exons and may help define exons (reviewed in [12]). Finally, some epigenetic marks on histones are associated with

regions of actively transcribed mRNAs, and a subset of these marks correlate with patterns of splicing and alternative splicing (reviewed in [13]). I next discuss (i) the major findings regarding proteins that regulate splicing and (ii) experiments suggesting links between transcription rates, nucleosome position, histone marks, and regulated alternative splicing.

Splicing Factors

SR proteins: SR proteins are a conserved class of eukaryotic proteins with known roles in the regulation of alternative splicing first isolated in mammalian cell culture (reviewed in [14-24]). Canonical mammalian SR proteins were defined by two characteristics: 1) an ability to complement and rescue a mammalian cell fraction (cytoplasmic S100 extracts) lacking the ability to splice reporter pre-mRNAs and 2) being recognized by monoclonal antibody 104 (mAb104) [19,25-30]. Cytoplasmic S100 extracts are derived from a fractionation protocol that removes SR proteins, but supplementation with either nuclear extracts or with purified SR proteins restores splicing competence to cytoplasmic S100 extracts.

Canonical SR proteins have similar protein structures. SR proteins contain one or two N-terminal RNA-recognition motifs (RRMs) and multiple C-terminal arginine and serine dipeptides (RS domain) [1,19,22,26], after which they are named. The RRMs of SR proteins are thought to recognize specific sequences of pre-mRNAs and, by doing so, to define exons. The RS domains are important for protein-protein interactions between SR proteins and other proteins required for splicing.

SR proteins play roles in splice site selection. Increasing concentrations of SR proteins in *in vitro* splicing assays promote use of proximal or downstream splice sites in

reporter pre-mRNAs (Figure 1.2B). All canonical SR proteins rescue splicing of S100 fractions with reporter pre-mRNAs, but some SR proteins promote splicing more efficiently than others. These results suggest both redundancy and specificity of SR proteins for pre-mRNAs.

The RRM domains of SR proteins interact with pre-mRNAs by their association with short degenerate motifs known as exonic splicing enhancers (ESEs) and intronic splicing enhancers (ISEs). Systematic evolution of ligands by exponential enrichment (SELEX) [31] experiments using mammalian [16] or *Drosophila* [32] SR proteins demonstrate that specific ESEs and ISEs are bound by individual SR proteins. Because of the degeneracy of ESE and ISE sequences, multiple SR proteins can be redundant and bind multiple ESE or ISE sequences (reviewed in [17]). Due to the degeneracy of ESEs, all eukaryotic exons likely contain ESEs of varying strength, although the degree of conservation of ESEs is uncertain based on the limited number of available studies. Studies of ESE distribution (defined by ESE sequences discovered through SELEX) in human mRNAs suggest enrichment of ESEs in both constitutive and alternative exons as compared to introns [33]. The strength of an ESE and the affinity of SR proteins for specific ESE sequences may determine the strength of 5' and 3' splice sites of alternative exons.

SR proteins are thought to be involved in both constitutive and alternative splicing. A general model has emerged in which SR proteins interact with ESEs in both constitutive and alternatively spliced exons. Once associated with pre-mRNAs, SR proteins interact with other proteins through their RS domains. Specifically, RS domains can interact with U2AF35 (U2 auxiliary factor subunit of 35 kD) at the 3' splice site to recruit the U2 small nuclear

ribonuclear protein (snRNP) to the branchpoint. RS domains can also interact with U1-70K that recruits U1 snRNP to the 5' splice site [34-36]. Recruitment of U2AF35 and U1 snRNPs marks exons to be spliced together in mature mRNAs (Figure 1.1B). Some alternative exons may not always be spliced into a mature message due to differences in the strengths of ESEs and the specificities for different SR proteins.

RS domains are found in a large number of proteins related to SR proteins, and many of them function in alternative splicing. Proteins containing RS domains are broadly categorized as SR-like. Well-characterized examples include members of the *Drosophila* sex-determination pathway such as TRANSFORMER and TRANSFORMER 2, the U2 auxiliary factors U2AF35 and U2AF65, U1-70K, and others (reviewed in [1,14]).

Canonical SR proteins have names based both upon their molecular weights and a numbered nomenclature. SR proteins were first purified from HeLa nuclear extracts and named based upon purified SR protein molecular weight. Seven proteins comprise the canonical SR family. A new nomenclature has renamed SR proteins as follows, with names in parentheses indicating the old SR nomenclature: SRSF1 (ASF/SF2), SRSF2 (SC35), SRSF3 (SRp20), SRSF4 (SRp75), SRSF5 (SRp40), SRSF6 (SRp55), and SRSF7 (9G8) [22,37] (see Table 1.1 for orthologous SR protein names in *C. elegans* and other species).

SR proteins are conserved in *C. elegans, Drosophila, Arabidopsis,* and mammals, but few examples are found in yeast. SR proteins are greatly expanded in plants (a new nomenclature for plant SR proteins has also been proposed, described in [38]). *Arabidopsis* SR proteins have been expanded to nearly 20 members [39] that fall into classes similar to mammalian canonical SR proteins, but contain multiple SR proteins within most classes.

Schizosaccharomyces pombe contains two SR proteins, SRp1 and SRp2. *Schizosacchromyces cerevisiae* has three SR-like proteins, Npl3, Hrb1, and Gbp2 [22,40].

C. elegans possesses seven *SR* genes, named *rsp-1* through *rsp-7* [41]. RNAi experiments in *C. elegans* directed against *rsp* family members suggest these genes are largely nonessential [41,42]. With the exception of *rsp-3,* RNAi knockdown of a single *rsp* results in no visible phenotype. RNAi knockdown of *rsp-3* results in late embryonic lethality. *Rsp-3* is the orthologue of human *SRSF1* (*ASF/SF2*) and has been shown to be an essential gene in chicken cell lines [43]. Defects are evident in worms treated with RNAi against two or more *rsp* genes. Double RNAi knockdown of *rsp-1* and *rsp-6* results in sterility and vulval defects; knockdown of *rsp-1* and *rsp-5* results in slower motility; knockdowns of [*rsp-2* and *rsp-4*], [*rsp-4* and *rsp-5*], or [*rsp-2* and *rsp-5*] show no phenotype. A triple knockdown of *rsp-2, rsp-4,* and *rsp-6* shows variable phenotypes, including slow growth, vulval defects, and Dpy phenotypes. Combined RNAi knockdown targeting *rsp-1, rsp-2, rsp-4, rsp-5,* and *rsp-6* genes is lethal. These observations are almost entirely consistent with my observations of *rsp* deletion alleles (described in Chapter 3). However, alleles of *rsp-6* used in this thesis are sterile as homozygotes. These RNAi experiments suggest *rsp* genes in worms are largely redundant.

Extensive alternative splicing of *SR* mRNAs is observed in worms, flies, plants, and mammalian species [44]. *SR* mRNAs in worms, plants, and mammals are not only alternatively spliced, but can be spliced to include PTCs in their mature mRNA transcripts [39,45-50]. *SR* splicing patterns change dramatically in *Arabidopsis* [39,46,51] and *C. elegans* during physiological stress [52,53]. The changes in splicing increase the proportion

of mature mRNAs that are nonfunctional due to incorporation of PTCs. PTC-containing mRNAs are degraded by NMD (see below), and increases in the proportion of PTC-containing mRNA suggest that regulating alternative splicing of *SR* mRNA is a means of reducing the expression of functional SR splicing factors.

SR proteins are also regulated by phosphorylation and by their intracellular localization in response to physiological stress. Such changes in SR protein phosphorylation and localization impact splicing of substrate mRNAs ([54] and reviewed in [55-57]). These observations suggest that SR proteins regulate diverse sets of transcripts, and the mRNAs that encode functional SR proteins can be regulated by changes in splicing patterns, phosphorylation, and localization.

hnRNPs: Heterogeneous nuclear ribonucleoproteins (hnRNPs) comprise a second class of conserved splicing regulators in eukaryotes (reviewed in [1,22,23,58-60]). hnRNPs were identified as proteins that associate with heterogeneous nuclear RNA (hnRNA) [61,62]. Immunoprecipitation of hnRNP family members identifies a large complex of copurifying proteins [62-64], including members of the hnRNP A, B, and C groups. Such proteins comprise a canonical set of hnRNP proteins that has been expanded to include more than 20 members (hnRNP A – hnRNP U) [58,59,64].

Immunodepletion of hnRNP from cell extracts suggested roles for hnRNPs in splicing. Mammalian cell extracts that are competent for *in vitro* splicing of reporter pre-mRNAs are splicing defective following immunodepletion of hnRNP proteins [59,62,65,66]. Analysis of hnRNP A1 suggested that it functions in splice site selection. In splicing assays similar to

those described for SR proteins, greater concentrations of hnRNP A1 in *in vitro* splicing reactions affect the use of distal or upstream 5' splice sites (Figure 1.2B). The effects of hnRNP A1 on splice site selection are the opposite of those caused by the previously described SR proteins, and hnRNP A1 effects can be inhibited by increased concentrations of the SR protein SRSF1 (also known as ASF/SF2) [9,58,67]. Thus, SR proteins can inhibit the effects of hnRNP A1 in splice site choice.

hnRNP proteins are structurally similar to each other and contain at least one RRM, with the exception of hnRNP U [60]. Some hnRNPs contain a 'K homology' (KH) domain instead of canonical RRMs, but many hnRNPs contain several RRMs of different RRM subclasses. As with *SR* mRNAs, *hnRNP* mRNAs are alternatively spliced, including alternative splicing that introduces PTCs to *hnRNP* mRNAs. SELEX experiments identified binding motifs for hnRNPs that may function in regulating alternative splicing [59,60,68]. hnRNP proteins have also been implicated in cellular roles other than splicing, including telomere maintenance, DNA repair, and chromatin remodeling (Reviewed in [60]).

Transcription and alternative splicing patterns

Most RNAs are spliced co-transcriptionally. Two models have emerged to explain how co-transcriptional splicing can affect patterns of alternative splicing: the "kinetic model" and the "recruitment model" (reviewed in [10,69]). The kinetic model posits that when the rate of transcription elongation by RNA Polymerase II (Pol II) is high, less time is available for alternative splice sites to be recognized [23,70,71]. Thus, faster rates of transcription elongation yield fewer instances of alternative splicing. Conversely, when the rate of Pol II

elongation is low over the length of a gene, splicing factors are more likely to recognize alternative (perhaps weaker) splice sites, which yields altered patterns of splicing (Figure 1.3 A). The recruitment model posits that splicing factors are brought to sites of transcription through interactions with the COOH-terminal domain (CTD) of Pol II during transcription. The abundance of splicing factors and the affinity of splicing factors for interactions with Pol II may affect the rate at which splicing factors are recruited to splice sites, thus influencing the patterns of splicing.

Support for the kinetic model comes from *in vivo* experiments with mutant Pol II polymerases that exhibit altered patterns of splicing that correlate with slower rates of Pol II elongation. For example, Pol II mutants (termed C4 mutants) with reduced rates of elongation show increased inclusion of the "EDI" (extra domain I) exon of human *FIBRONECTIN* mRNA [11,69,72].

Support for the recruitment model comes from both *in vitro* and *in vivo* experiments. Immunopurification of Pol II followed by mass spectrometry demonstrates wild-type Pol II interacts *in vitro* with snRNPs and SR proteins through the CTD of Pol II [73]. These results led to experiments showing that the presence of SR proteins are required at the very beginning of a coupled *in vitro* transcription and splicing reaction for both efficient transcription and efficient splicing [73]. If SR proteins are added after the initiation of transcription, both transcription and splicing are inefficient. These results demonstrate that SR proteins function during transcription *in vitro*, presumably to splice pre-mRNAs co-transcriptionally, and are recruited to sites of splicing through interactions with the CTD of Pol II. *In vivo* experiments with the *FIBRONECTIN* "EDI" exon also demonstrate support for the recruitment model [74].

Overexpression of SRSF3 (SRp20) inhibits inclusion of the EDI exon; silencing of SRSF3 by RNAi increases the inclusion of the EDI exon. Silencing of SRSF3 by RNAi in Pol II mutants that lack a CTD has no effect on EDI inclusion. The effect of SRSF3 on EDI inclusion through the CTD of Pol II is independent of the rate of transcription. The experiments described above provide evidence that supports both the kinetic and recruitment model of co-transcriptional splicing. However, in principle, both models are able to explain co-transcriptional splicing.

Correlations between nucleosome positioning, histone modifications, and splicing

The position of nucleosomes correlates with the position of exons in multiple eukaryotic species (Reviewed in [12,13,75-77]). Nucleosomes are octamers comprised of four histones, H2A, H2B, H3, and H4 [78]. The positions of nucleosomes on chromatin can be inferred by sensitivity of DNA to cleavage by micrococcal nuclease, which makes single-stranded nicks in nucleosome-associated DNA but double-stranded cuts in linker DNA between nucleosomes. Thus, nucleosome positions can be mapped by the pattern of micrococcal nuclease cleavage sites. Global analyses of human, *Drosophila*, and *C. elegans* chromatin with micrococcal nuclease demonstrates that nucleosomes are disproportionately enriched in exons having weak splice sites and associated with exon inclusion [79-83]. Similar findings using a computational model to assign a nucleosome occupancy score (NOScore) surrounding splice junctions suggest nucleosome occupancy is higher in exons, including constitutive, cassette, and exons with alternative 5' or 3' splice sites [84].

Nucleosome and exon positions correlate, as do their average sizes in a strikingly simple and elegant manner. 147 base pairs of DNA are wrapped in a nucleosome. The average size of a metazoan exon is approximately 145 base pairs [76]. The similar size of nucleosomes and exons, and the enrichment of nucleosomes at exons have led to models suggesting nucleosomes may help distinguish exons from introns.

Many studies have investigated the global patterns of epigenetic marks in eukaryotic chromatin. Many of these marks affect expression at genetic loci, but some are associated with exons and alternatively spliced exons (Reviewed in [76,77,85-88]). H3K36me3 is a histone mark associated with actively transcribed DNA, particularly at the 3' end of genes [89]. H3K36me3 marks are enriched at internal exons and particularly enriched at constitutive exons of actively transcribed genes [82,90-93]. H3K36me3 marks also correlate with increased nucleosome occupancy [83]. Similar correlations between alternatively spliced exons and H3K36me3 marks in human chromatin have been observed [93].

H3K36me3 marks correlate with tissue-specific alternative splicing of the human *fibroblast growth factor receptor 2* (*FGFR2*) [94,95]. *FGFR2* exons IIIb and IIIc are mutually exclusive cassette exons and are alternatively spliced. Exon IIIb is found in prostate epithelial cells (PNT2), while exon IIIc is found in human mesenchymal stem cells (hMSC). H3K36me3 marks are enriched in hMSC cells and correlate with exon IIIc inclusion. SET2 is an H3K36me3 methyltransferase, and overexpression of SET2 causes global increases in H3K36me3 marks on *FGFR2*. Overexpression of SET2 further reduces exon IIIb inclusion in hMSC cells. Silencing of SET2 via RNAi increases exon IIIb inclusion in hMSC cells. An adaptor protein, MRG15, associates with H3K36me3 marks and recruits the splicing regulator

polypyrimidine tract binding protein 1 (PTBP1) to regulate the splicing of IIIb and IIIc exons (Figure 1.1 B). Similar correlations between H3K36me3 marks and regulated splicing were observed for *TPM1*, *TPM2*, and *PKM2* mRNAs. These results suggest that splicing factors can be recruited to sites of alternative splicing by proteins that recognize histone modifications. Recruited splicing factors can then repress inclusion of an alternative exon.

Non-SR and non-hnRNP proteins can act in *trans* to influence splicing, likely by influencing Pol II transcription rates or DNA modifications. The DNA binding protein CCCTC-binding factor (CTCF) binds exon 5 of the human *CD45* gene and influences whether exon 5 is included in mature *CD45* mRNA transcripts, likely by influencing the rate of Pol II elongation [96,97]. *In vitro* nuclear run-on experiments demonstrate that CTCF binding to CCCTC sites in a target gene slows the rate of RNA Pol II elongation (Figure 1.3 C). This process is likely inhibited by 5-methylcytosine DNA methylation to block binding sites of CTCF as methylated DNA immunoprecipitation (MedIP) shows a strong inverse correlation between methylation sites and CTCF binding sites at *CD45* exon 5. siRNAs that direct an increase in H3K9me2 marks in target genes may cause an increase in alternative exon inclusion [98]. One model proposes that acetylated H3K9 marks allow rapid transcription of alternatively spliced genes and siRNAs direct an increase in both H3K9me2 and H3K27me3 marks that slow elongation of Pol II and allow for increased rates of alternative splicing.

Nucleosome occupancy and DNA modifications that correlate with exon position suggest alternative splicing may be regulated in part by nucleosome position and histone

modifications. However, the importance of nucleosome position and histone modifications in regulated alternative splicing is still a matter of debate.

Developmentally regulated and tissue-specific alternative splicing

The abundance of alternative splicing in higher eukaryotes compared to that of lower eukaryotes suggests that the genomic coding potential expanded considerably during evolution. Alternative splicing patterns differ during development and in specific tissue-types. Such patterns have been observed in mammalian tissues, *Drosophila,* worms, and in *Arabidopsis* (reviewed in [5])

Alternative splicing patterns change during development in *Drosophila* and *C. elegans*. Global analyses of 30 growth stages and conditions in *Drosophila* [6] and nearly 20 in *C. elegans* [8] document extensive changes in the patterns of alternative splicing of both species during their life cycles. In *Drosophila*, approximately 60% of alternatively spliced mRNAs change patterns of splicing, and in *C. elegans* approximately 30% of alternatively spliced mRNAs change patterns throughout development.

Many tissue-specific, alternative splicing regulators have been defined (reviewed in [23]). For example, NOVA1 and NOVA2 act to regulate brain specific splicing patterns [99-101]. Other regulators include a neuronal paralog of PTB (also known as hnRNP I), nPTB. nPTB and PTB show different patterns of regulation dependent upon neuronal differentiation, and PTB levels are low when nPTB levels are high in differentiated neurons [102-105].

Alternative splicing is extensive among eukaryotes, but precisely how tightly controlled regulation occurs during development and in different tissues is poorly understood.

A combination of previously described models is likely. Differences in the quantity of specific splicing factors, the state of chromatin, and the rates of Pol II elongation in different stages of development and/or tissue types all possibly influence tissue-specific and developmentally regulated alternative splicing.

I have thus far described instances and the prevalence of alternative splicing, as well as the molecular mechanism of alternative splicing. In the next section, I will discuss a particularly interesting class of alternatively spliced mRNAs, those spliced to include PTCs. PTC-containing mRNAs are substrates of the NMD pathway. I first describe the NMD pathway and the current understanding of the mechanism of NMD. I then describe mRNAs that are substrates of NMD, including those that are alternatively spliced mRNAs to contain PTCs, and the regulation of alternative splicing to generate such mRNAs.

Nonsense-mediated mRNA decay (NMD)

PTC-containing mRNAs are actively and rapidly degraded in all tested eukaryotes, including yeast, worms, flies, mammals, and plants (Reviewed in [106-118]). Studies of NMD focus on two primary questions: how does NMD work and why does NMD exist? I first describe the factors required for NMD and the molecular mechanism of NMD, which is conserved in eukaryotes. I then discuss the *in vivo* function of NMD.

NMD Factors

NMD was first described in yeast and worms. In yeast, up-frameshift 1 (*UPF1*) was isolated in a forward genetic screen for suppressors of the *his4-38* frameshift mutation [119].

The *his4-38* allele results in a +1 frameshift causing a PTC within the *HIS4* reading frame. *Upf* mutants block NMD and allow synthesis of HIS4 protein in *his4-38* mutants resulting in a His^+ phenotype. *UPF1, UPF2,* and *UPF3* have since been shown to also be required for NMD in yeast [120,121]. *C. elegans smg* (suppressor with morphogenic effects on genitalia) mutants suppress the paralysis phenotype of *unc-54(r293)* and were recovered in forward genetic screens [122,123]. The *unc-54(r293)* allele deletes a portion of *unc-54* 3' UTR, including the polyadenylation and cleavage site, and expresses an *unc-54* mRNA with an abnormally long 3' UTR causing the "normal" stop codon to be perceived as premature. Thus, mRNAs of *unc-54(r293)* are substrates for NMD. *Smg* mutations suppress the paralysis phenotype of *r293* by eliminating NMD, which causes *unc-54(r293)* mRNA to be stable rather than unstable.

Seven *smg* genes (*smg-1* through *smg-7*) described in *C. elegans* and other eukaryotes are required for NMD [123-128]. Yeast contains three genes required for NMD, *UPF1, UPF2,* and *UPF3,* which are orthologues of *C. elegans smg-2, smg-3,* and *smg-4*, respectively [119-121]. *Smg* genes are conserved in higher eukaryotes and are named after either their *smg* or *UPF* gene founders (reviewed in [117]) with the exception of *Drosophila,* which lacks an orthologue of *smg-7* [129]. All of the *C. elegans smg* genes are required for NMD, but none is essential for viability. Orthologues of certain *smg* genes are essential in a number of higher eukaryotes, including mammals, Zebrafish, plants and *Drosophila* (reviewed in [117]).

Recent experiments identified additional NMD factors in both *C. elegans* and human cell culture. Genome-wide RNAi screens revealed two essential genes, *smgl-1* and *smgl-2* (*smg-lethal*), that are involved in NMD in *C. elegans* [130]. RNAi knockdown of the human

orthologues, *SMGL-1* and *SMGL-2*, inhibit NMD in humans [130]. Analysis of proteins that copurify with human SMG-1 identified two additional genes, *SMG-8* and *SMG-9*, that may exhibit weak NMD defects in *C. elegans* and human cells when silenced by RNAi [131].

NMD Mechanism

SMG-2 is the central regulator of NMD. SMG-2 undergoes cycles of phosphorylation and dephosphorylation that are required for NMD [125]. SMG-1 is a phosphatidylinositol 3-kinase (PIKK) family member and phosphorylates SMG-2 [124,132]. SMG-3 and SMG-4 are also required for SMG-2 phosphorylation, though their precise role in SMG-2 phosphorylation is unknown. SMG-5 interacts with PP2A, a phosphatase, to dephosphorylate SMG-2 [126]. SMG-6 and SMG-7 are also required to efficiently dephosphorylate SMG-2. SMG-7 and SMG-5 interact to bring PP2A to SMG-2. Despite extensive research, the biochemical roles for SMG-2 phosphorylation are still unknown.

SMG-2 preferentially associates with PTC-containing mRNAs in *C. elegans* [133], yeast [134], and humans [135]. Thus, SMG-2 marks PTC-containing mRNAs before such mRNAs are degraded. SMG-2 interacts with translation release factors 1 and 3 (eRF1 and eRF3) [136,137] and is likely brought to terminating ribosomes in a complex with eRF1,3. Subsequent to marking, downstream factors act to degrade PTC-containing transcripts (reviewed in [138]). SMG-6 acts as an endonuclease to cleave PTC-containing mRNAs in *Drosophila* [139] and mammals [140], and exonucleases degrade cleaved mRNA. How PTC-containing mRNAs are destabilized in *C. elegans* is unknown.

Translation is required for NMD in all organisms tested (reviewed in [141]). Interactions between translation termination factors and SMG-2 may recruit SMG-2 to mRNAs undergoing translation. The position of translation termination within an mRNA can affect whether NMD occurs. Translation termination factors interact with poly(A)-binding (PABP) proteins, and the distance between these factors is important to discriminate translation that terminates at a normal stop or a premature stop [142-145]. If translation terminates at a normal stop, SMG-2 and translation termination factors interact with factors bound at the 3' end of an mRNA, including PABP. However, if translation terminates far from the 3' end of the mRNA, SMG-2 and 3' factors are separated. Thus, distance between SMG-2 and 3' factors may play a role in discriminating improper translation termination and signaling for NMD to degrade PTC-containing transcripts. (Figure 1.4A).

NMD is enhanced in vertebrates and plants by the deposition of the exon junction complex (EJC) at splice junctions [117]. The EJC is a multi-protein complex composed of factors involved in mRNA splicing, export, and translation. UPF2 and UPF3 (SMG-3 and SMG-4 orthologues) are bound to the EJC and help recruit UPF1 (SMG-2) to the EJC and sites of translation termination. Though the EJC can enhance NMD, a 3' EJC is not a requirement to elicit NMD, as the presence of a long 3' UTR and PTCs within the last exon (downstream of an EJC) can both elicit NMD [129].

Biological Significance of NMD

A long-standing question in the NMD field concerns the biological functions of NMD. Initial hypotheses suggested NMD acts as a proofreading mechanism to remove errors that

arise through normal biological processes that introduce PTCs into mature mRNAs [123,146]. During the life of an mRNA, such errors could happen at many points. Errors in transcription could introduce a PTC at random within an mRNA. Errors in splicing could lead to frame-shifts in mature mRNAs by not including exons or retaining introns. While such mRNAs are likely NMD substrates, mRNAs containing PTCs as a result of errors in transcription or splicing associated with introduction of PTCs have not been found. Errors in these processes are rare, and therefore repeatedly demonstrating such errors is difficult. However, errors in DNA replication and mutagens can introduce nonsense mutations into genes that are transcribed into PTC-containing mRNAs.

Genome-wide transcriptome analyses by both microarray and high-throughput sequencing suggest NMD actively degrades a very large number of eukaryotic mRNAs [134,147-155]. Global analyses in yeast, worms, flies, mammals, and plants demonstrate approximately 5% to 20% of mRNAs increase in abundance when NMD is blocked. Two studies have distinguished between mRNAs that are direct substrates of NMD and those that are not [134,152]. Direct substrates are defined as mRNAs that Upf1p (SMG-2) associates with or mRNAs whose half-lives are affected by NMD. These analyses indicate that nearly half of all mRNAs whose abundance increases when NMD is blocked are substrates of NMD. The remaining half of mRNAs whose abundance increases when NMD is blocked presumably represents the indirect effects of the absence of NMD.

Global analyses of mRNA abundance in NMD defective mutants suggest that NMD substrates fall into classes based upon the way in which a PTC is introduced to substrate mRNAs. Large classes of mature mRNAs that are substrates for NMD include pseudogenes,

mRNAs with 5' upstream open reading frames (uORFS), mRNAs subject to "leaky scanning" (ribosomes that initiate translation downstream of the proper start codon), and mRNAs containing PTCs as a result of alternative splicing. The types of PTC-containing mRNAs (uORFs, pseudogenes, etc.) appear common across eukaryotic species, but many mRNAs that are substrates for NMD in one species may not be an NMD substrate in a different species. An obvious exception, however, are *rsp* and *rpl* (described in the next section) mRNAs that are NMD substrates in several species.

Pseudogenes are nonfunctional genes that have acquired mutations over time. The abundance of mRNAs of *C. elegans* expressed pseudogenes increase in NMD-deficient mutants [156]. Expressed pseudogenes are substrates of NMD because they contain many PTCs and translation frameshifts. Global analyses of mRNA abundance in NMD mutants suggest expression of pseudogenes is common, but NMD acts to mitigate their expression. Nearly 25% of yeast pseudogenes are upregulated when NMD is blocked [155].

Intron retention, uORFs, and leaky scanning appear to be the predominant types of PTC-containing mRNAs elevated in NMD defective yeast [155]. In humans, uORFs are also a significant class of NMD substrates, as well as intron-containing 3' UTRs, transposons, and alternatively spliced mRNAs [154]. Tiling arrays and high-throughput sequencing in *C. elegans* suggest that most NMD substrates result from errors in splicing, but also from 5' uORFs and mRNAs with long 3' UTRs [149]. Alternative splicing and long 3' UTRs are also common among *Drosophila* NMD substrates [147].

mRNAs alternatively spliced to include PTCs in their mature mRNAs are a particularly intriguing class of NMD substrates. Many alternative splicing events appear to be

tightly regulated and are mechanisms eukaryotes use to regulate gene expression. These events are described in the next section.

Regulated unproductive splicing and translation (RUST)

The mRNAs of a remarkably high proportion of eukaryotic genes are alternatively spliced [2,5-8]. A seemingly counter-intuitive outcome of some alternative splicing is the generation of mature mRNAs that contain PTCs. As previously described, PTC-containing mRNAs are rapidly degraded in eukaryotes. Why then do many eukaryotes make such mRNAs if the immediate outcome is their destruction?

"Regulated unproductive splicing and translation" (RUST) is a model that posits alternative splicing to generate PTC-containing mature mRNAs is a mechanism to regulate gene expression [47,157,158] (Figure 1.4 B) Alternatively spliced mRNAs that contain PTCs are commonly referred to as resulting from “unproductive” (as opposed to "productive") splicing events. I describe unproductively spliced mRNAs as “PTC” or "PTC-containing" mRNAs in following chapters.

RUST suggests that alternatively spliced mRNAs are produced as a means of post-transcriptional regulation of gene expression. Early reports of RUST suggested RUST allows proteins encoded by alternatively spliced mRNAs to influence the splicing patterns of their own pre-mRNAs. An excess of protein can shift splicing to increase the proportion of PTC-containing mRNA, and a deficit of protein can alter splicing to decrease the proportion of PTC-containing mRNA. Small changes in the ratios of alternative isoforms may have big

impacts on gene expression. Thus, RUST can influence splicing patterns to affect fine changes in gene expression.

Numerous examples of RUST have been described. The Anderson lab described alternative splicing of four *C. elegans rpl* (ribosomal protein large subunit) mRNAs to include or exclude PTCs in their mature mRNAs by using alternative 5' or 3' splice sites [159]. Overexpression of RPL-12, for example, increases the proportion of *rpl-12(PTC)* mature mRNAs, suggesting that RPL-12 protein directly or indirectly directs the splicing of its own mRNA. Intriguingly, mRNAs of the mammalian orthologues of *rpl* genes are alternatively spliced in a similar manner. For example, rat RPL-3 regulates splicing of its own mRNA in a manner similar to that of *C. elegans* RPL-12 [160].

Proteins that regulate alternative splicing appear to be enriched among the many documented examples of RUST. Examples include mammalian SR family members *SRSF2*, *SRSF3*, and *SRSF4*, whose patterns of splicing are regulated by the proteins they encode [47,50,161]. RUST of *SR* mRNAs is likely conserved. The regions adjacent to *SR* alternative splicing events are ultraconserved in mouse and human, suggesting that splicing of *SR* mRNA is regulated by RUST in both species [47-49]. RUST also regulates splicing of hnRNP mRNAs, including *hnRNP L*, *hnRNP LL* [162], *PTB* (also referred to as *hnRNP I*) [163], and *nPTB*, a neuronal *PTB* paralog [102,104].

The prevalence of RUST-mediated regulation in eukaryotes is a matter of ongoing debate. The earliest estimates of RUST suggested that nearly a third of all alternatively spliced human mRNAs may be spliced to introduce PTCs [157,158]. These estimates were suggested to be an overestimation by later studies that used microarrays to interrogate known

alternative splicing events in mammalian cell cultures where NMD was inhibited [164]. The use of microarrays, however, precludes discovery of previously unknown alternative splicing events. Considering recent increases in estimates of alternative splicing of human genes, it is likely that previous microarray analyses missed many alternative splicing events that might introduce PTCs. Despite the uncertainty of the overall prevalence of RUST, splicing regulators appear to be a class of mRNAs disproportionately regulated by RUST [47,151,165].

A surprising recent finding indicates that Pol II occupancy, mRNA expression, and alternative splicing patterns are linked to NMD in an elegant manner [52]. Although such regulation may be a different mechanism of regulation than the traditional RUST models, it may represent an additional layer of regulation. When Pol II elongation rates are perturbed, Pol II occupancy increases at introns flanking alternative exons in a large number of mRNAs. Increases in Pol II occupancy at introns flanking alternative exons correlate with increased alternative exon inclusion and decreased mRNA expression. A subset of these mRNAs are alternatively spliced to include PTCs, thus generating substrates of NMD. Alternatively spliced mRNAs that are substrates of NMD are overrepresented among alternatively spliced exons correlated with increased Pol II occupancy and decreased mRNA expression. The authors suggest a model wherein decreased rates of transcription allow for weaker splice sites to be utilized, consistent with the kinetic model of alternative splicing. As a consequence of slower transcription, less mRNA is made and the splicing patterns of these mRNAs ensure an even further reduction in their abundance by shuttling mature mRNA to the NMD pathway.

Introduction to future chapters

The following chapters of this dissertation focus on describing alternative splicing of the *C. elegans rsp* mRNAs. Chapter 2 describes initial identification of *rsp* mRNAs as substrates of the NMD pathway and characterizations of *rsp* splicing patterns that subject *rsp* mRNAs for degradation by NMD. I observed *rsp* mRNA expression and the splicing patterns of several *rsp* mRNAs are dramatically regulated during *C. elegans* development. Chapter 3 describes my experiments to understand the roles that individual RSP proteins play in regulating splicing of both their own and each other's pre-mRNAs. Such regulated events are examples of RUST, and I observed patterns of both feedback autoregulation and cross-regulation. In Chapter 4, I summarize the results and conclusions of this thesis and speculate on possible mechanisms by which *rsp* splicing patterns are governed.

Figure 1.1. Alternative splicing of eukaryotic mRNAs.

A) Types of alternative splicing patterns observed in eukaryotes. Lines represent introns, boxes represent exons, and dashed lines indicate splicing patterns. Yellow exons are constitutive exons and red exons represent alternative exons. B) Model of how SR and hnRNP proteins regulate splicing patterns by defining exons. The blue box represents an alternatively spliced exon and the lines on either end represent introns upstream and downstream. SR proteins (yellow ovals) associate with exonic splicing enhancer sequences (ESEs) within alternative exons and recruit U2AF subunit of 35 kD (orange oval) to the 3' splice site and U1 snRNP (green oval) to the 5' splice site. hnRNPs (red oval) associate with exonic splicing suppressor (ESSs) sequences within the alternative exon to inhibit inclusion of alternative exons. SR proteins can inhibit hnRNP association with ESSs. This figure is adapted from [87].

Figure 1.1

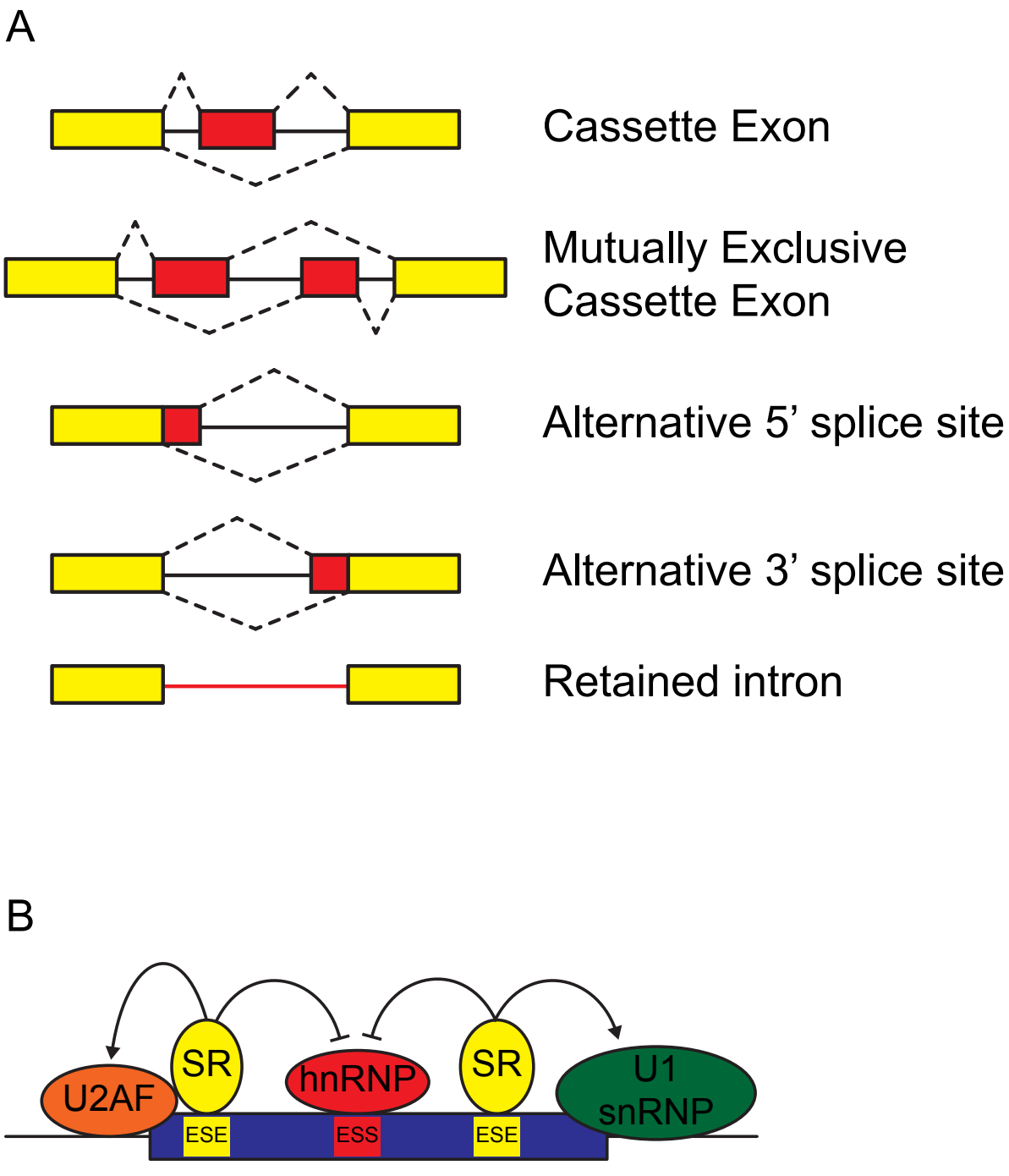

A
Cassette Exon
Mutually Exclusive
Cassette Exon
Alternative 5' splice site
Alternative 3' splice site
Retained intron
B
U2AF
SR
hnRNP
SR
U1
snRNP
ESE
ESS
ESE

Figure 1.2. SR and hnRNP proteins play roles in 5' splice site selection.

Early studies of SR and hnRNP function in splice site selection used radiolabeled pre-mRNAs with two or more 5' splice sites (A). Yellow boxes represent constitutively spliced exons. Proximal splice sites (red box) are defined as the furthest downstream splice site and distal splice sites (blue box) are defined as the furthest upstream splice sites. Red lines represent proximal splicing patterns and blue lines represent distal splicing patterns. B) SR and hnRNP proteins affect the selection of 5' splice sites. In *in vitro* splicing reactions, pre-mRNAs with multiple 5' splice sites (shown in A) are added to extracts capable of splicing reporter mRNAs. Purified SR or hnRNP proteins are added to these extracts and the products of these reactions are resolved in acrylamide gels. Increasing concentrations of SR or hnRNP proteins generally have opposite effects on splice site selection. Increased concentrations of SR proteins promote the use of proximal splice sites to yield mature mRNAs containing the proximal 5' splice site (middle row). Increased concentrations of hnRNP proteins in general have the opposite effect and promote the use of distal splice sites to yield mature mRNAs with the distal 5' splice site (bottom row).

Figure 1.2

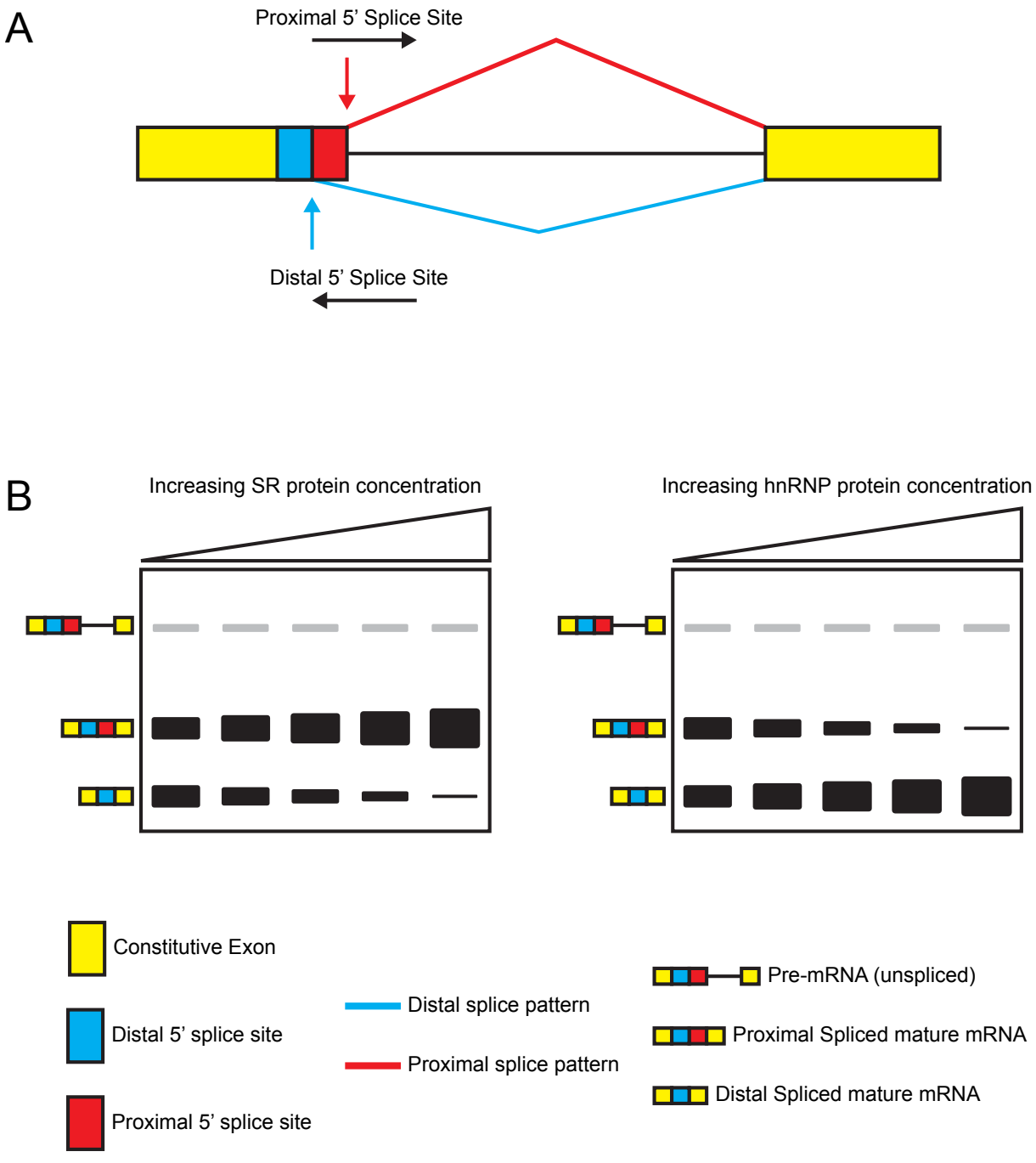

A
Proximal 5' Splice Site
Distal 5' Splice Site
B
Increasing SR protein concentration
Increasing hnRNP protein concentration
Constitutive Exon
Distal 5' splice site
Proximal 5' splice site
Distal splice pattern
Proximal splice pattern
Pre-mRNA (unspliced)
Proximal Spliced mature mRNA
Distal Spliced mature mRNA

Table 1.1. SR protein nomenclature across eukaryotes.

SR proteins were first described in mammalian cell culture where they were named by molecular weight (Historical mammalian SR protein names). A new nomenclature has renamed mammalian SR proteins as SR splicing factor 1-12 (SRSF1 – SRSF12). Orthologues of SR proteins in multiple eukaryotic species are shown. This chart has been adapted from [22,37,41]. Expanded classes of *Arabidopsis* SR proteins are described in [39,51] and a new plant SR nomenclature is detailed in [38]. SR proteins are found in at least 27 eukaryotes and are described in [44].

Table 1.1

New Mammalian SR Protein	Historical Mammalian SR Protein	*C. elegans*	*Drosophila*	*S. pombe*	*S. cerevisiae*	*Arabidopsis*
SRSF1	ASF, SF2, SRp30a	RSP-3	SF2			RS31a, SR1, SRp34a, SR34b,
SRSF2	SC35, PR264	RSP-4, RSP-5	SC35	Srp1		SCL28, SCL30, SCL30a,
SRSF3	SRp20	RSP-6	Rbp1, Rbp1-like, Rsf1			RSZ21, RSz22, RSZ22a,
SRSF4	SRp75	RSP-1		Srp2	NPL3	RSP41, RSP31, RSP40
SRSF5	SRp40, HRS	RSP-2				
SRSF6	SRp55, B52		B52			
SRSF7	9G8	RSP-6	Xl6			
SRSF8	SRp46					
SRSF9	SRp30c					
SRSF10	TASR1, SRp38, SRp40					
SRSF11	p54, SRp54	RSP-7	Srp54			
SRSF12	SRp35					

Figure 1.3. Models of the effects of transcription rates and histone modifications on alternative splicing patterns.

A) The kinetic model of alternative splicing proposes that alternative exons (red) are less likely to be included in mature mRNAs when RNA polymerase II (purple oval) transcribes quickly along the length of a gene (top). When RNA Pol II rates of transcription are slow (bottom), alternative exons are more likely to be included in mature mRNAs along with constitutive exons (yellow). Orange cylinders represent nucleosomes wrapped by DNA. B) *FGFR2* splicing is affected by H3K36me3 marks that affect inclusion of exon IIIb or IIIc in *FGFR2* mature mRNA. H3K36me3 marks (star) modify nucleosomes (orange cylinder). MRG15 (teal oval) associates with H3K36me3 marks and recruits PTB1 (brown rectangle) to inhibit inclusion of exon IIIb in *FGFR2* mature mRNA. C) CTCF is a DNA-binding protein (green oval) that can inhibit the rate of Pol II transcription and slowed Pol II rates of transcription result in inclusion of exon 5 in the mature *CD45* mRNA transcript. B and C are adapted from [87].

Figure 1.3

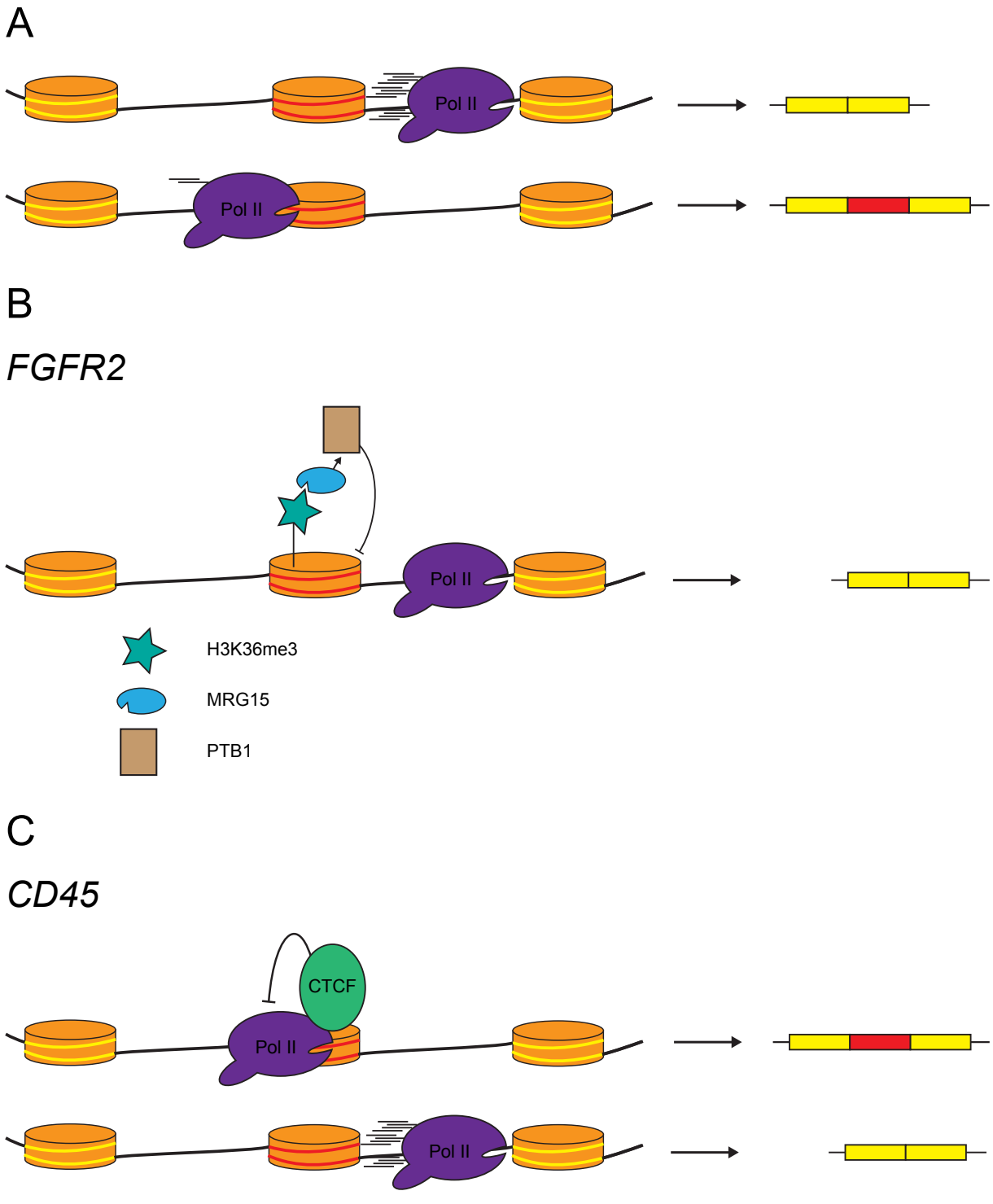
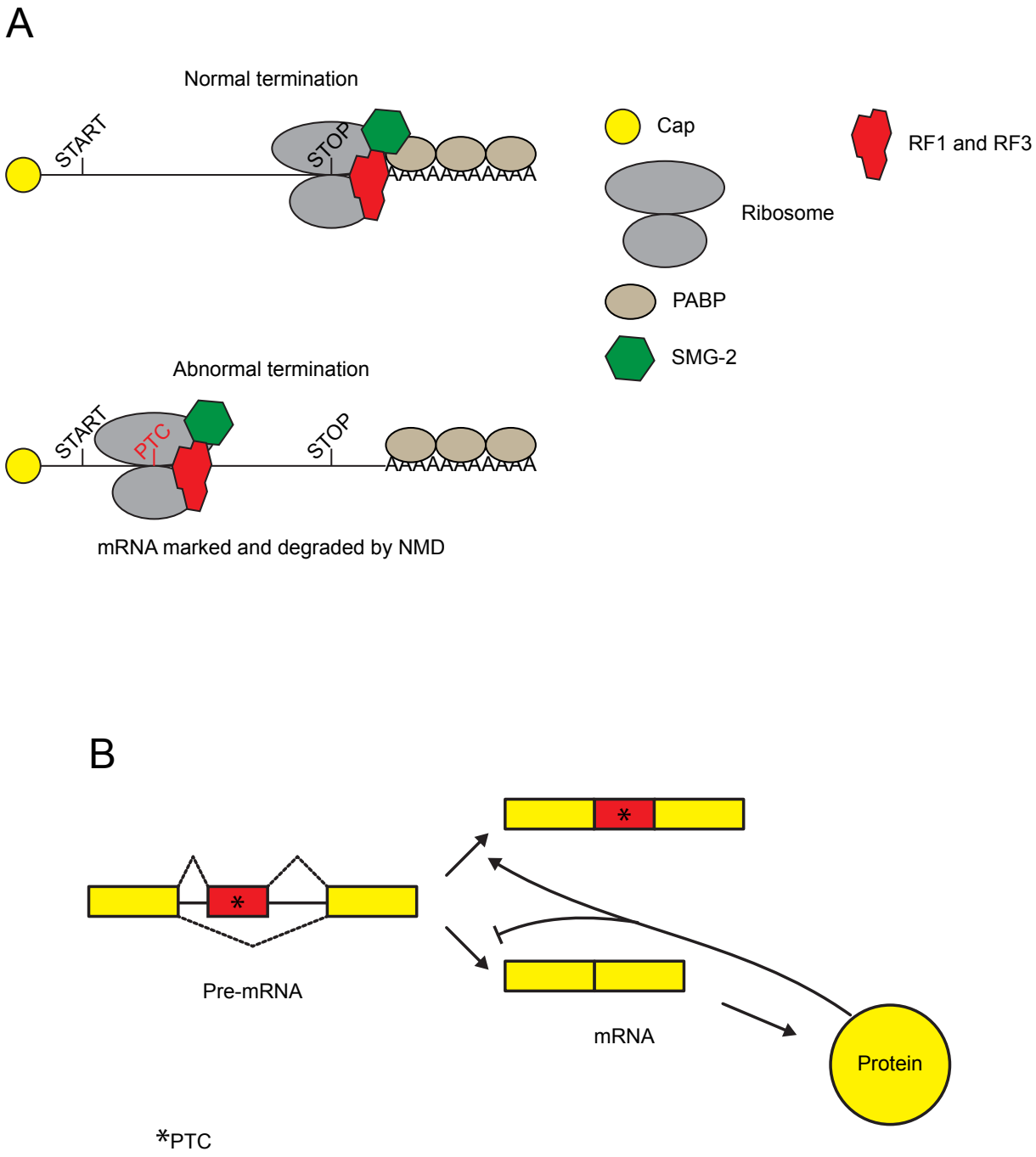

A
Pol II
Pol II
B
FGFR2
Pol II
H3K36me3
MRG15
PTB1
C
CD45
CTCF
Pol II
Pol II

Figure 1.4. Models of translation termination and Regulated Unproductive Splicing and Translation (RUST).

A) During normal translation termination, SMG-2 (green hexagon) associates with poly(A)-binding protein (PABP) (tan oval) through interactions with translation release factors (red shape) that bring SMG-2 in proximity with PABP. When SMG-2 is far from PABP translation termination is abnormal, triggering NMD. B) mRNAs translated into functional proteins may directly or indirectly affect the splicing patterns of their own mRNAs. If protein (yellow oval) levels are high, proteins can influence the splicing pattern of their own pre-mRNAs to increase the proportion of mature mRNAs containing PTCs (red exon) to maintain protein homeostasis.

Figure 1.4


A
Normal termination
START
STOP
AAAAAAAAAAA
Cap
RF1 and RF3
Ribosome
PABP
SMG-2
Abnormal termination
START
PTC
STOP
AAAAAAAAAAA
mRNA marked and degraded by NMD
B
*
Pre-mRNA
mRNA
Protein
*PTC

Chapter 2: Developmental Regulation Of *C. elegans rsp* mRNAs

Abstract

Significant proportions of eukaryotic transcriptomes are alternatively spliced [2,5]. A central function of alternative splicing is to increase protein diversity arising from pre-mRNAs [2]. A second and significant function is to generate transcripts that contain premature termination codons (PTCs) within their open reading frames [47]. The presence of PTCs in alternatively spliced mature mRNAs destabilizes those mRNAs via the nonsense-mediated mRNA decay (NMD) pathway. Thus, alternatively spliced PTC-containing mRNAs are effectively destroyed.

I observed that alternative splicing of *C. elegans rsp* mRNAs introduces PTCs into mature *rsp* mRNAs. Such alternative splicing is likely conserved in a closely related nematode species, *C. briggsae*, as the nucleotides adjacent to *rsp* splicing events are conserved between *C. elegans* and *C. briggsae*. Orthologous *rsp* mRNAs of other eukaryotic species are similarly spliced to introduce PTCs in their mature mRNAs. The conservation of sequence and similarity of alternative splicing patterns suggest that the decision to splice *rsp* pre-mRNAs in a manner that contains or does not contain a PTC is regulated.

This chapter outlines experiments I conducted to investigate the patterns of *rsp* mRNA splicing and to determine if and how such events are regulated during *C. elegans* development. I find that six of the seven *rsp* pre-mRNAs are alternatively spliced in a manner that yields one or more *rsp* mature mRNAs containing PTCs in their open reading frames. Expression of six alternatively spliced *C. elegans rsp* genes exhibit similar patterns of expression during development. *Rsp* mRNAs are of highest abundance in embryos and are

dramatically downregulated in early and subsequent larval stages. The proportion of PTC-containing [*rsp(PTC)*] and PTC-free [*rsp(+)*] mature mRNAs correlates with the abundance of *rsp(+)* mRNA for four of the six alternatively spliced *rsp* mRNAs. Increased *rsp(+)* mRNA expression, such as observed in the embryo, correlates with a decreased proportion of *rsp(PTC)* mRNAs, and decreased *rsp(+)* mRNA expression, such as observed in larval stages, correlates with an increased proportion of *rsp(PTC)* mRNAs. I investigated whether changes in *rsp(+)* mRNAs correlate with changes in RSP protein abundance. I observed changes in the protein abundance of RSP-6 that roughly parallel changes in *rsp-6(+)* mRNA abundance.

Introduction

Alternative splicing is common in eukaryotes. Recent estimates suggest that mRNAs of 90-100% of human genes are alternatively spliced [2,4,5]. Messenger RNAs of lower eukaryotes are also alternatively spliced, approximately 60%, 25%, and 42% in *Drosophila*, *C. elegans*, and *Arabidopsis*, respectively [5-8]. Though lower than rates in humans, alternative splicing in lower eukaryotes is a major feature of gene expression.

Alternative splicing generates multiple mature mRNA isoforms from the same pre-mRNA, and a major role of alternative splicing is to increase the diversity of the proteome [2,5,166]. For example, alternative splicing in humans generates tissue-specific protein isoforms [5,167]. The sequences of tissue-specific, alternatively spliced exons are highly conserved, indicating that both protein function and the pattern of splicing are selected during evolution [167].

A second major role of alternative splicing is the generation of PTC-containing mRNAs. For example, PTC-containing mRNAs arise from alternative splicing through inclusion or exclusion of exons and through alternative 5' or 3' splice sites. PTC-containing mRNAs are substrates for NMD and are degraded. PTC-containing mRNAs arising from alternative splicing are observed in mammals, plants, and worms (reviewed in [47]). Many of these splicing events are regulated and are described in more detail below.

Alternative splicing is highly regulated, and such regulation occurs throughout eukaryotic development. Alternative splicing of regulators of the *Drosophila* sex-determination pathway is a well-studied example (reviewed in [168]). Expression of *Sex-lethal* (*Sxl*) is activated in early development in females and establishes a positive feedback loop that regulates splicing of its own pre-mRNA. SXL protein binds its own pre-mRNA in females and promotes exclusion of the male-specific exon 3, which in males generates a nonfunctional male-specific SXL protein. Female-specific SXL establishes a cascade of downstream, regulated, gender-specific, alternative splicing of gender-specific mRNAs such as *transformer, doublesex, fruitless*, and others [1,169].

Large numbers of mRNAs are alternatively spliced in *C. elegans*, and many of these splicing events are regulated during development [8,170,171]. Sequencing of the transcriptome during *C. elegans* development identified 349 alternatively spliced mRNAs whose isoforms change significantly between at least two developmental stages [8]. Such developmentally regulated alternative splicing represents approximately 30% of all alternative splicing. Comparable studies in *Drosophila* suggest approximately 66% of alternatively spliced mRNAs are developmentally regulated [6].

Alternative splicing that introduces PTCs to mature mRNAs are well-documented (reviewed in [47] and detailed in chapter 1). Such mRNAs have been described as being "unproductively" spliced, but I refer to them in this thesis as "PTC-containing" mRNAs. A general model proposes that *regulated* alternative splicing generates PTC-containing mRNA as a mechanism to regulate protein expression. This model has been termed regulated unproductive splicing and translation (RUST) and is described in further detail in Chapter 3 [157,158].

Factors that regulate alternative splicing, such as SR and hnRNP proteins, are themselves alternatively spliced in human, *Drosophila, Arabidopsis, C. elegans*, and many other eukaryotes [22,44]. SR proteins are serine/arginine-rich proteins that act as positive regulators of alternative splicing and promote inclusion of alternatively spliced exons into mature mRNA transcripts (reviewed in [22]). Heterogeneous nuclear ribonuclear proteins (hnRNPs) generally act in an opposing manner, promoting exclusion of exons in mature mRNAs. SR proteins can antagonize hnRNP interactions with pre-mRNAs.

Generation of PTC-containing *SR* mRNAs is common amongst eukaryotes. In *C. elegans* (observations described in further detail in the following section) one or two additional PTC-containing isoforms are observed for *rsp* mRNAs, orthologous to *SR* mRNAs. In *Arabidopsis,* nearly two-thirds of *SR* mRNAs are spliced to include PTCs [39], and approximately half of these are substrates for NMD [44,172]. In humans and mice all *SR* mRNAs are alternatively spliced to introduce PTCs either in their open reading frame or in their 3' UTR and become substrates for NMD [48,49]. Generation of PTC-containing *SR* mRNAs through alternative splicing therefore suggests conserved functions. I observed

alternative splicing of *C. elegans rsp* mRNAs and hypothesized these events are regulated. This chapter describes initial characterizations of *C. elegans rsp* splicing patterns and investigates their regulation during *C. elegans* development.

Results

Rsp mRNAs are alternatively spliced and degraded by NMD in *C. elegans*

Prior work in the Anderson lab utilized microarray analysis to identify *C. elegans* mRNAs whose abundance increases in *smg(-)* mutants (Dave Markwardt, unpublished observations). Among mRNAs increased in *smg(-)* mutants were several members of the *rsp* gene family, including *rsp-2, rsp-6,* and *rsp-7*. Northern blots of *rsp* mRNAs in N2 (wild type) and *smg-1(-)* mutants suggested that their increased expression in *smg(-)* mutants might be due to alternatively spliced *rsp* mRNAs that are only detected in *smg(-)* mutants (Dave Markwardt, unpublished observations). Work in another lab also demonstrated that the abundance of *C. elegans rsp-4* and *rsp-6* are increased in *smg(-)* mutants [45]. *Rsp* genes in other species are similarly elevated in NMD-defective cells [47-50,172]. Alternative splicing to generate transcripts with PTCs degraded by NMD is a commonly observed phenomenon, therefore I hypothesized *rsp* mRNAs are similarly alternatively spliced.

I tested if *rsp* mRNAs are alternatively spliced in *C. elegans* by RT-PCR. I designed primers to walk along the length of each *rsp* transcript and performed RT-PCR in wild type and *smg-2(-)* mutants to determine if additional transcripts are stabilized in *smg-2(-)* mutants. I chose to use *smg-2(-)* mutants because SMG-2 is the central regulator of NMD, and *smg-2(-)* mutants have a strong NMD-defective phenotype [133]. In wild-type L4-stage worms, I

observe a single predominant band for *rsp-1* through *rsp-7* (Figure 2.1A). In *smg-2(-)* mutants, I observe the same band as in wild type and, with the exception of *rsp-3*, one or more additional bands for each *rsp* gene (Figure 2.1A).

Because my initial results suggested the additional *rsp* transcripts that accumulate in *smg(-)* mutants are the result of alternative splicing, I sequenced cDNAs of all transcripts detected in both wild type and *smg-2(-)* mutants. Transcripts detected in *smg-2(-)* mutants but not in wild type result from alternative splicing in which an additional exon is included in the mature *rsp* mRNA, and this inclusion introduces a PTC (Figure 2.1B). I term mRNAs that express full-length canonical RSP proteins as "*rsp(+),*" and those that contain one or more PTCs as "*rsp(PTC)*" in the following text. Both *rsp-1* and *rsp-6* are spliced in at least three ways (see figure 2.1B): (i) without an additional exon [e.g. *rsp-1(+)*] , (ii) with only one additional exon [e.g. *rsp-1(PTCshort)*], or (iii) with two additional exons [e.g. *rsp-1(PTClong*)].

Alternatively spliced *rsp* mRNAs share several characteristics. First, all alternative splicing events use canonical *C. elegans* 5'-GU and 3'-AG splice sites. Second, all alternative-splicing events occur in what had previously been annotated as the largest intron, with the exception of *rsp-7*. Third, the alternative splicing events introduce PTCs by inclusion of an additional exon. Interestingly, orthologous *rsp* mRNAs alternatively spliced in other eukaryotes exhibit similar patterns of exon inclusion, and SR proteins are often associated with promoting inclusion of additional exons in mature mRNA messages. From these results, I conclude six of the seven *C. elegans rsp* mRNAs are alternatively spliced and

the manner of alternative splicing introduces a PTC. Thus, these transcripts are likely targets of the NMD pathway.

The abundance of alternatively spliced *rsp* mRNAs increases in multiple *smg* mutants

Because SMG-2 has non-NMD functions, including roles in stauffen-mediated mRNA decay [173], I decided to test whether alternatively spliced *rsp* mRNAs increase in abundance in additional *smg* mutants. Increased abundance in multiple *smg* mutants is strong evidence that PTC-containing mRNAs are substrates for NMD. Like in *smg-2(-)* mutants, the abundance of *rsp(PTC)* mRNAs increases in *smg-1(-)*, *smg-3(-)* and *smg-5(-)* mutants (Figure 2.1A). These data show *rsp(PTC)* transcripts are degraded by NMD and I conclude that alternatively spliced *rsp* mRNAs are substrates of NMD.

Conservation of *rsp* alternative splicing

Regions surrounding alternatively spliced exons of *SR* mRNAs are conserved in higher eukaryotes [48,49], and I suspected regions surrounding *rsp* alternative exons might also be conserved among worms. I investigated whether the regions surrounding alternatively spliced exons are conserved using the VISTA genome browser [174]. I visualized the nucleotide conservation of *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7* between *C. elegans* and *C. briggsae*, two closely related nematode species (Figure 2.2). The VISTA genome browser plots the average percent conservation between species within a specified window. For example, the VISTA browser calculates the overall percent conservation in 100 bp surrounding base pair X. It repeats this for base pair Y, Z, and so on. The calculated percent

conservations at base pairs X, Y, Z, etc. are then displayed and can be interpreted. Regions with higher than 70% conservation are considered highly conserved.

In general, for the *rsp* alternatively spliced regions, the conservation is higher than surrounding regions. For alternatively spliced exons of *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7*, the conservation is higher than the introns surrounding them. The conservation of alternatively spliced *rsp* exons suggests shared functions between *C. elegans* and *C. briggsae*. I hypothesize the conservation of alternative exon regions is due to alternative splicing of *rsp* exons in both species.

qRT-PCR Assays

Because multiple *rsp* transcripts arise via alternative splicing, I hypothesized these events are regulated. To investigate regulation of *rsp* alternative splicing, I decided to use qRT-PCR to measure *rsp* splicing patterns and mRNA abundance. Since alternatively spliced *rsp* mRNAs are very similar, *rsp* primers were carefully designed to detect specific *rsp* transcripts.

Rsp qRT-PCR primers were designed to be transcript-specific and to meet the following criteria: (1) The primers discriminate between *rsp(PTC)* and *rsp(+)* mRNAs; (2) the primers span exon-exon junctions, such that only mature mRNA is amplified; (3) the primers yield products from only one transcript; (4) the primers do not amplify "primer dimers"; (5) the sequence of amplified products matches a single *rsp* mRNA; and (6) a melt curve after qRT-PCR shows accumulation of a single amplified PCR product. As outlined in Figure 2.3, primers used to quantify individual *rsp(PTC)* mRNAs anneal partly within the

PTC-containing exon and do not amplify *rsp(+)* mRNA. Primers used to quantify *rsp(+)* mRNAs span an exon-exon junction unique to *rsp(+)* mRNA and do not amplify *rsp(PTC)* transcripts. Additional primers were designed to amplify total *rsp* mRNA. Such primers span exon-exon junctions common to all mRNAs [both *rsp(+)* and *rsp(PTC)*] of each *rsp* gene. Primers that met all the above criteria were used to quantify individual mRNAs by qRT-PCR. In measurements that follow, the short and long forms of *rsp-1(PTC)* and *rsp-6(PTC)* were summed to generate a single quantification for *rsp-1(PTC)* and *rsp-6(PTC)* mRNA. In all experiments, *rsp-1(PTC-short)* and *rsp-1(PTC-long)* show similar patterns of expression, as do *rsp-6(PTC-short)* and *rsp-6(PTC-long)*.

The abundance of *rsp* mRNAs and the proportions of PTC-containing vs. PTC-free mRNAs are regulated during development

I investigated whether expression of *rsp* genes is regulated during development and whether the proportion of PTC-free vs. PTC-containing splicing changes during development. Regulation of alternative splicing is widespread among eukaryotes, including during development. Thus, I hypothesized that *rsp* alternative splicing might be regulated during development to influence *rsp* gene expression. RSP protein expression is almost certainly influenced in part by mRNA expression, and therefore RSP expression might be regulated in part by regulation of *rsp* alternative splicing. Because NMD degrades transcripts with PTCs, *rsp* mRNAs might be spliced to include PTCs when an excess of *rsp* mRNA is present. Such a mechanism would allow for a fine-tuning of *rsp* mRNA abundance, such that appropriate levels of *rsp(+)* mRNA can be maintained or adjusted.

I quantified *rsp* mRNA during *C. elegans* development in synchronized populations of embryos, the four larval stages, and adults. I independently quantified expression of *rsp(+)* and *rsp(total)* in both wild type and *smg-2(-)* mutants, as well as *rsp(PTC)* in *smg-2(-)* mutants. *Rsp* mRNA expression was quantified relative to expression of *eft-3*, a normalization control. *Eft-3* mRNA levels were consistent across developmental stages in qRT-PCR experiments and microarray analysis of *eft-3* during *C. elegans* development show mostly stable expression throughout development [175]. I then calculated the proportion of *rsp(PTC)* and *rsp(+)* mRNAs at each stage. NMD efficiently degrades *rsp(PTC)* mRNAs in wild type, making their abundance difficult to quantify reproducibly. I therefore quantified expression of *rsp(PTC)* mRNAs only in *smg-2(-)* mutants, in which NMD does not occur. I assume the *smg-2* mutation only affects whether *rsp(PTC)* transcripts are degraded and not the splicing of *rsp(PTC)* transcripts since *rsp(PTC)* transcripts accumulate in multiple *smg(-)* mutants.

Expression of *rsp-1(+), rsp-2(+), rsp-4(+), rsp-5(+), rsp-6(+),* and *rsp-7(+)* mRNAs decrease 2.5- to 5-fold in wild type L1 larvae compared to embryos (see Figure 2.4A). Expression remains low through larval development [*rsp-2(+)*, *rsp-4(+)*, and *rsp-7(+)*] or increases slightly [*rsp-1(+)*, *rsp-5(+),* and *rsp-6(+)*]. Expression of *rsp(+)* in adults is approximately half that of embryos, with the exception of *rsp-5(+)*, whose expression in adults is approximately equal to that of embryos. However, measurements of *rsp-5(+)* mRNA were more variable than other *rsps*. *Rsp(PTC)* mRNAs are a very small fraction of total *rsp* mRNAs in wild-type strains. Thus, both *rsp(+)* and *rsp(total)* mRNAs show similar levels of expression at all tested stages.

Expression of *rsp-1(+), rsp-2(+), rsp-4(+), rsp-5(+), rsp-6(+),* and *rsp-7(+)* mRNAs in *smg-2(-)* mutants exhibits a similar developmental pattern as wild type (see Figure 2.4B). Expression decreases approximately 3-fold [*rsp-5(+)* and *rsp-7(+)*] to 10-fold [*rsp-4(+)* and *rsp-6(+)*] in L1 larvae compared to embryos and remains low during larval development. The modest increase of *rsp-1(+)*, *rsp-2(+)*, *rsp-4(+)*, and *rsp-6(+)* mRNA in late larval and adult stages is less pronounced in *smg-2(-)* mutants than in wild type.

Expression of *rsp-1(PTC), rsp-2(PTC), rsp-4(PTC), rsp-5(PTC),* and *rsp-6(PTC)* in *smg-2(-)* mutants generally parallels those of *rsp(+)* mRNAs. The decrease of *rsp(PTC)* mRNAs that occur in L1 larvae relative to embryos, however, is not as great as for *rsp(+)* mRNAs in those same samples (2- to 5-fold decrease of *rsp(PTC)* mRNAs compared to 3- to 10-fold decrease of *rsp(+)* mRNAs; see Figure 2.4B). Expression of *rsp-7(PTC)* decreases steadily during larval development.

I initially hypothesized that splicing of *rsp* pre-mRNAs is regulated to modulate expression of *rsp* mRNAs and proteins. Figure 2.5 quantifies the expression of *rsp(PTC)* mRNA as a percentage of the total *rsp* mRNA for each gene [*rsp(PTC)*/sum of *rsp(PTC)* and *rsp(+)* mRNAs]. Splicing that produces the highest proportion of *rsp(PTC)* mRNA (95%) occurs for *rsp-6* at the L1 stage, and splicing that produces the lowest proportion of *rsp(PTC)* mRNA (10%) occurs for *rsp-4* at the L4 and adult stages. In general, the proportion of PTC-containing mRNA is lowest in embryos, the stage where *rsp(+)* mRNA abundance is at its highest. Similarly, the proportion of PTC-containing mRNA is generally highest in mid-larval stages, where *rsp(+)* mRNA abundance is at its lowest. The proportion of PTC-

containing mRNA for *rsp-2* and *rsp-6* remains high from L1 through adult stages, while that of other *rsp* genes declines in adults, especially *rsp-4* and *rsp-7*.

The changing proportions of *rsp(+)* vs. *rsp(PTC)* mRNAs that I observed during development in *smg-2(-)* mutants explains, in part, the changing relative expression of *rsp(+)* mRNA that I observed in wild type. Alternative splicing, however, cannot fully explain the magnitude of all changes in *rsp(+)* mRNA. For example, *rsp-2(+)* mRNA decreases approximately 5-fold in wild type L1 animals compared to the embryo (Figure 2.4A). The fraction of PTC-containing *rsp-2* mRNA increases from 55% in the embryo to 80% in the L1. A modest increase in PTC-containing mRNA cannot account for the large decrease in *rsp-2(+)* mRNA. Clearly, other regulatory mechanisms, such as decreased *rsp-2* transcription or increased *rsp-2(+)* mRNA degradation, also contribute to the abundance of *rsp-2(+)* mRNA. Similar arguments apply to *rsp-1*, *-4*, *-5*, *-6*, and *-7*.

Confirmation of *rsp* mRNA regulation by northern blots

The qRT-PCR data presented above are very consistent, but to confirm the quantitative results using an independent method, I performed northern blots and quantified the abundance of *rsp* mRNAs in wild type and *smg-2(-)* mutants. Although the size difference between *rsp(PTC)* and *rsp(+)* mRNAs is approximately 100 bp or less, I found that they could be reliably resolved on 3% agarose gels containing DMSO and glyoxal and electrophoresed for extended lengths of time. I hybridized northern blots with antisense, single-stranded, RNA probes generated by *in-vitro* transcription of *rsp* templates. The probes

hybridize to both *rsp(PTC)* and *rsp(+)* mRNAs (Figure 2.6A). An antisense RNA-probe of *eft-3* was used to normalize *rsp* mRNAs.

Patterns of *rsp* mRNA expression measured with northern blots are very similar to those measured by qRT-PCR (Figure 2.6 A-C). Using independent biological samples, I measured *rsp-1, rsp-2, rsp-4, rsp-5,* and *rsp-6* mRNAs in embryo, L1, and L2-stage wild-type and *smg-2(-)* samples. I chose to only observe the first three stages since the most dramatic changes in *rsp* mRNA abundance occur between the embryo and L1 transition.

Northern blots are unable to distinguish *rsp-7(PTC)* and *rsp-7(+)* transcripts for two reasons. First, the size difference is only approximately 80 bp between the *rsp-7(PTC)* and *rsp-7(+)* isoforms. Second, *rsp-7* transcripts are longer than other *rsp* mRNAs I tested (1359 bp for *rsp-7(+)* and 1435 for *rsp-7(PTC)*). I was unable to resolve these size differences even running gels as long as 24 hours. I used two separate *rsp-7* probes for northern blotting, and both probes detected one band of the size expected for *rsp-7* mRNA (data not shown). Neither probe detects two bands. Semi-quantitative RT-PCR show the *rsp-7* splicing pattern in the first three developmental stages matches the patterns I observed with qRT-PCR (Figure 2.7) as described in the previous section. In semi-quantitative experiments, the proportion of *rsp-7(PTC)* mRNA changes in a similar pattern as changes observed in qRT-PCR assays. Though I was able to measure the splicing pattern of *rsp-7* with semi-quantitative PCR, this assay is insufficient to measure individual *rsp* mRNA transcripts throughout development. I conclude based on two independent lines of evidence: qRT-PCR and northern blots, *rsp* mRNA expression patterns are developmentally regulated, and the most dramatic regulations occur between the embryo and L1 transition.

Developmental regulation of RSP-6 protein

Changes in the relative mRNA abundance of *rsp(+)* mRNAs during development suggested there might be corresponding changes in RSP proteins. To test the expression patterns of RSP proteins during development, I attempted to prepare affinity-purified anti-RSP antibodies against RSP-2, RSP-4, RSP-5, and RSP-6. Only the anti-RSP-6 sera is able to reproducibly and specifically detect an RSP protein. Anti-RSP-6 is specific for RSP-6 protein and detects an approximately 20 kDa protein present in *rsp-6(+)* worms, but absent in two separate deletion alleles of *rsp-6(-)* worms *ok798* and *tm367* (data not shown for *tm367*) (Figure 2.8).

To test if RSP-6 is developmentally regulated, I grew synchronized wild-type and *smg-2(-)* worms and collected embryo, four larval stages, and adult-stage samples. In both wild type and *smg-2(-)* mutants, RSP-6 is expressed highly in the embryo, L1, L2, and adults (Figure 2.9). RSP-6 decreases dramatically between L2 and L3 stages and remains low in both L3 and L4 animals. I conclude that expression of RSP-6 is developmentally regulated. Expression of RSP-6 roughly parallels abundance of *rsp-6(+)* mRNA in wild type (Figure 2.4A), with one notable exception. The relative expression of *rsp-6(+)* mRNA declines 2-3 fold in L1s, but the relative expression of RSP-6 remains approximately constant. The data suggest either that (i) RSP-6 expressed in embryos has a longer half-life (persisting through the L1 stage); or (ii) *rsp-6* mRNA is translated more efficiently in L1 larval stages. I have not, however, measured relative rates of RSP-6 degradation or *rsp-6* translation. Thus, contributions of regulated translation or RSP-6 stability to overall expression of RSP-6 during

development are unknown. Expression of RSP-6 roughly parallels abundance of *rsp-6(+)* mRNA in *smg-2(-)* mutants with the same exception noted above.

Discussion

Alternatively spliced *rsp* mRNAs are stabilized in NMD-defective mutants

Messenger RNAs of six of the seven *C. elegans rsp* genes are alternatively spliced to include PTCs in their mature mRNAs. The abundance of *rsp(PTC)* mRNAs and of total *rsp* mRNA is greatly increased in each of four tested *smg(-)* mutants (Figure 2.1), and I assume that the elevated abundance reflects instability of *rsp(PTC)* mRNAs in NMD-competent cells. I therefore conclude that *rsp(PTC)* mRNAs are substrates of NMD and are actively degraded in wild type.

Conservation of alternative splicing generating PTC isoforms

Alternative splicing to generate PTC-containing *SR* mRNAs is observed in many eukaryotes including mice, humans, plants, and worms [47]. The regions surrounding alternatively spliced exons are very highly conserved ("ultraconserved," see [48,49]) in mammals, indicating that splicing to generate PTC-containing mRNA is selected in evolution. The regions surrounding alternatively spliced exons in *C. elegans* are also highly conserved among nematode species. Conservation of sequence adjacent to alternatively spliced regions suggests alternative-splicing events play important regulatory functions. Conservation of sequences also suggest regulatory functions are evolutionarily conserved. *SR* PTC-containing mRNAs are also generated in *Arabidopsis.* The *Arabidopsis* genome encodes 19 *SR* genes,

and mRNAs of 15 of them are alternatively spliced, resulting in 95 distinct *SR* transcripts [39]. Of these 95 transcripts, approximately 60 contain PTCs and nearly half of PTC-containing *SR* transcripts are substrates of NMD [172]. Alternative splicing of *SR* mRNAs is clearly expanded in plants, but, like in mammalian and lower eukaryotic species, appears to be conserved. Alternative splicing of *SR* mRNAs in *Arabidopsis* appears to be regulated by stress, in particular temperature and hormones and such stresses give rise to several PTC-containing *SR* mRNAs [39]. Splicing patterns of *SR* mRNAs throughout eukaryotes indicate generation of PTC-containing mRNAs is a common event and suggests a shared mechanism of *SR* gene regulation.

Other splicing regulators are similarly spliced to include PTCs in a subset of mRNAs. hnRNPs are generally negative regulators of alternative splicing that act in a manner opposite to that of SR proteins, which promote alternative exon inclusion. A *C. elegans* hnRNP mRNA, *hrpf-1* is spliced to generate PTC-containing *hrpf-1* mRNAs [referred hereafter as *hrpf-1(PTC)*] [171]. If exon 5 of *hrpf-1* is excluded, a frameshift results in a PTC. Inclusion of *hrpf-1* exon 5 results in PTC-free mature mRNA encoding functional HRPF-1 protein. Patterns of alternative *hrpf-1* splicing are conserved in *C. briggsae*. *hnRNP* mRNAs are also alternatively spliced in mammalian species to introduce PTCs similar to *C. elegans hrpf-1* by exclusion of an exon to introduce a PTC [104,162,163].

Generation of PTC-containing mRNA by splicing similar to that of *rsp* genes occurs for mRNAs that do not encode splicing factors. For example, mRNA of *rpl-12*, a *C. elegans* ribosomal protein gene, is alternatively spliced in a manner similar to *rsp* mRNAs. The

sequence of *rpl-12* surrounding alternatively spliced exons is also highly conserved in *C. elegans* and *C. briggsae*, as is the mammalian orthologue of *rpl-12* [159,160].

Thus, regulated alternative splicing to produce mRNAs that contain PTCs appears to be a common event. Conservation of nucleotide sequence between closely related and more distantly related eukaryotic species in regions adjacent to alternative splicing events suggest these regions have important functions that have been maintained throughout evolution.

Regulated expression and splicing of *rsp* mRNAs during development

I hypothesized that generation of PTC-containing *rsp* mRNAs might be a mechanism for regulating or fine-tuning expression of *rsp* genes, and this chapter describes that regulation during development. I quantified expression of *rsp(+)* mRNAs of six different *rsp* genes in *smg(+)* animals, as well as both *rsp(+)* and *rsp(PTC)* mRNAs of six different *rsp* genes in *smg(-)* animals. My data reveal the balance of PTC-free and PTC-containing splicing for each gene at each developmental stage.

Expression of *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7* mRNAs is developmentally regulated. The abundance of *rsp(PTC), rsp(+)*, and *rsp(total)* mRNAs of most *rsp* genes decrease 2- to 5-fold between embryo and L1/L2 stages, and the reduction of *rsp(+)* mRNAs is greater than that of *rsp(PTC)* mRNAs. The proportions of *rsp-2(PTC), rsp-4(PTC), rsp-6(PTC),* and *rsp-7(PTC)* mRNAs increase significantly during that period. The proportion of *rsp(PTC)* mRNA of these genes presumably also increases in wild-type animals, but, because NMD destabilizes *rsp(PTC)* mRNAs, I was unable to measure changing proportions of *rsp(PTC)* transcripts in wild type.

Because of the increase in the proportion of *rsp(PTC)* mRNA, I conclude that *rsp(PTC)* splicing is regulated during the embryo and L1 transition. I suggest that changes in the proportion of *rsp(PTC)* vs. *rsp(+)* mRNA is a mechanism that contributes to the overall regulation of *rsp* gene expression. The increase in the proportion of *rsp(PTC)* mRNA, however, is not sufficient to fully explain the 2.5- to 5-fold decrease in *rsp(total)* mRNA that occurs between embryo and L1 in wild type. The majority of this downregulation must occur by some other mechanism. Decreased abundance of *rsp(total)* mRNA could result from decreased rates of *rsp* transcription, decreased *rsp* mRNA half-life, or both.

Hrpf-1 alternative splicing is developmentally regulated similarly to *rsp* alternative splicing in *C. elegans* development. Between the embryo and L1 transition, the splicing pattern of *hrpf-1* changes. In L1 the proportion of mature *hrpf-1* spliced to include a PTC increases dramatically [171]. I hypothesize that *hrpf-1* splicing during development is regulated similarly to *rsp* mRNAs to downregulate *hrpf-1(total)* mRNA.

Despite the differences in the way *hrpf-1* and *rsp* mRNAs are spliced to include a PTC both *hrpf-1* and *rsp* mRNAs increase the proportion of PTC mRNA between embryo and L1, suggesting that less *hrpf-1(+)* and *rsp(+)* mRNA is needed in L1 animals. These results also suggest that regulated alternative splicing is a mechanism to regulate the level of (+) mRNA. Alternative splicing of *rsp* mRNAs may exemplify a widespread mechanism to regulate gene expression.

Rates of transcription might affect *rsp* splicing patterns during development

Rates of mRNA transcription can influence splicing of alternatively spliced mRNAs (reviewed in [10,71] and described in detail in chapter 1). Slower *in vitro* rates of RNA Polymerase II transcription, measured by Pol II occupancy on alternatively spliced mRNAs, correlate with increased rates of alternative splicing. Because most splicing occurs co-transcriptionally, alternatively spliced exons with weak splice sites are utilized more frequently when the rate of Pol II transcription decreases.

Global analysis of Pol II occupancy in mammalian cell cultures [52] show increases in alternative splicing of mRNAs that correlate with decreases in mRNA expression when Pol II occupancy decreases. A subset of these alternatively spliced mRNAs are spliced to include PTCs in their mature mRNAs, similar to *rsps* and *hrpf-1*. The subset of alternatively spliced mRNAs that contain PTCs is enriched among mRNAs whose alternative splicing patterns appear affected by Pol II occupancy.

The increase in the proportion of *rsp(PTC)* mRNA I observed during *C. elegans* development is not sufficient to account for the large decrease in *rsp(+)* mRNA. Therefore, changes in the rate of transcription that also affect patterns of alternative splicing might be a mechanism that explains both the downregulation of *rsp(+)* mRNA and the change in the *rsp(PTC)* vs. *rsp(+)* splicing patterns during *C. elegans* development (Figure 2.10A). However, changes in transcription alone may not affect *rsp* splicing patterns, but could still explain the downregulation of *rsp(+)* mRNA.

Indirect evidence supports a model of changes in transcription affecting rates of *rsp* splicing patterns. *C. elegans* experiments involving starvation demonstrate that Pol II

occupancy and alternative splicing are correlated. *Rsp-6, rsp-7,* and *hrpf-1* are among a group of mRNAs whose alternative splicing changes when animals are starved [52,53]. Starved L1 worms have increased stalling of Pol II at introns flanking the alternatively spliced exons, decreased mRNA expression, and altered splicing patterns for *rsp-6, rsp-7,* and *hrpf-1* mRNA isoforms. The proportion of *rsp-6(PTC), rsp-7(PTC),* and *hrpf-1(PTC)* increases in starved L1s, but upon refeeding, those proportions decrease.

The observations described above suggest downregulation of *rsp* and *hrpf-1* transcription between embryo and L1-stage worms could result in a coarse adjustment of *rsp* and *hrpf-1* total mRNAs. If Pol II occupancy and Pol II transcription rates do affect *rsp* and *hrpf-1* splicing patterns, decreased rates of transcription could coarsely turn down *rsp* and *hrpf-1* mRNA and in turn affect a fine adjustment of these mRNAs by changing the rates PTC isoforms are spliced. However, a coarse adjustment by changes in transcription could still explain the decrease in *rsp(+)* mRNA without affecting *rsp* and *hrpf-1* splicing patterns.

What might signal for changes in transcription rates of *rsps*? The transcription factors that regulate *SR* gene transcription are unknown, but chromatin states could influence transcription of *rsp* genes. Changes in the state of chromatin during development might influence the rate at which Pol II transcribes *rsp* pre-mRNAs. Such a model has been proposed for tissue-specific regulation of alternative splicing in mammalian species [5] and might be extended to developmental regulation of *rsp* splicing. H3K36me3 marks modify histones and have been observed to correlate with exon positions in worms [91]. These marks have been shown to affect tissue-specific splicing of exons of the human *FGFR2* transcript by indirectly recruiting splicing regulators that affect the splicing of *FGFR2* [94]. Such marks

might therefore be regulated at *rsp* loci, and could possibly change throughout development. Histone marks might influence the splicing patterns of *rsp* mRNAs by recruiting splicing factors, including RSP proteins, to alternatively spliced *rsp* exons.

The abundance of *rsp(+)* mRNAs in embryos might be maternal, and increased degradation rates might downregulate *rsp(+)* mRNAs during development

Regulating the rates of *rsp* and *hrpf-1* gene transcription is an appealing model for regulating alternative splicing, but other models are also possible. For example, *rsp* mRNAs may be maternally deposited. In *C. elegans,* transcription in the soma of early embryos begins at the four-cell stage, and after the 100-cell stage in the germline [176]. Maternally deposited *rsp* mRNAs might be stable and could be translated in early embryo development but might rapidly be degraded once zygotic transcription begins. If so, *rsp* mRNAs transcribed after zygotic transcription has begun would likely be in far less abundance than maternal mRNAs, especially if the bulk of maternal mRNA was degraded.

Whether abundant mRNAs in embryo are maternal or not, the rate of degradation could be the same. Other degradation pathways may be responsible for further decreasing *rsp* mRNA in L1-stage animals (see Figure 2.10B). Degradation could occur due to any number of known degradation pathways, such as changes in mRNA processing affecting the stability of *rsp* mRNAs, expression of endogenous siRNAs targeted against *rsp* mRNAs during development, and others.

Regulation of RSP-6 protein expression during development

The abundance of RSP-6 protein roughly parallels that of *rsp-6* mRNA in wild type (see Figure 2.4A and 2.9), but the correlation between mRNA and protein is not perfect. *Rsp-6(+)* mRNA decreases precipitously in L1 larvae, but RSP-6 does not. Perhaps RSP-6 has a longer half-life or *rsp-6* mRNA is translated more efficiently in L1 larvae. Further experiments would be required to clarify the relationship between *rsp-6(+)* mRNA abundance and RSP-6 protein abundance.

Although *rsp-6(+)* mRNA and RSP-6 protein expression do not correlate completely, I observed a consistent downregulation of both *rsp-6* mRNA and protein during development. The functional significance of reduced RSP-6 protein in L3 and L4 worms is unknown; such changes might well influence global patterns of alternative splicing. RSP-6 might regulate fewer alternatively spliced mRNAs in L3 and L4 stages. Identifying mRNAs whose splicing patterns are regulated by RSP-6 and other RSPs might provide valuable insights into why RSP-6 protein is downregulated.

Figure 2.1. *rsp* mRNAs are alternatively spliced and accumulate in NMD mutants.

A) RT-PCR of *rsp-1* through *rsp-7*. In wild type (N2), primarily only one transcript accumulates. In *smg*(-) mutants (defective for NMD), I observed additional *rsp* transcripts from *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7*. These additional mature *rsp* transcripts include a PTC within their open reading frame. I term PTC containing transcripts as *rsp(PTC)* and those lacking a PTC as *rsp(+)*. B) Maps of the alternative splicing events of each *rsp*. Boxes represent exons and lines represent introns. Red exons indicate an exon that if included in the mature transcript introduces a PTC. The position of the PTC is indicated in the mature transcript by an asterisk.

Figure 2.1

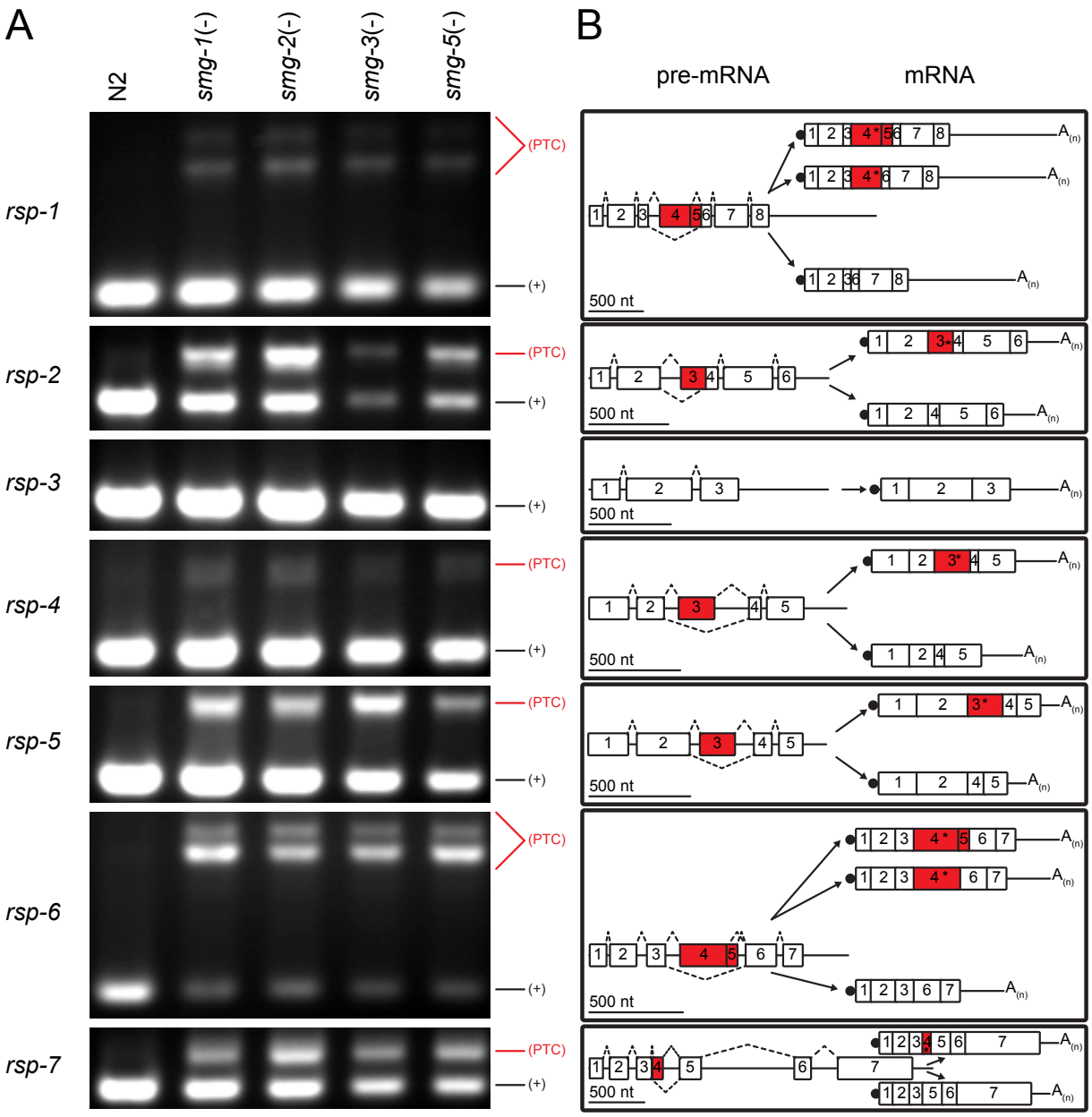

A
N2
smg-1(-)
smg-2(-)
smg-3(-)
smg-5(-)
rsp-1
rsp-2
rsp-3
rsp-4
rsp-5
rsp-6
rsp-7
(PTC)
(+)
B
pre-mRNA
mRNA
500 nt

Figure 2.2. VISTA genome browser shows alternatively spliced *rsp* mRNAs are conserved in regions surrounding alternative exons.

Alternatively spliced *rsps* are conserved between *C. elegans* and *C. briggsae*.

The VISTA genome browser plots the average percent conservation between the two species within a specified window. For example, the VISTA browser calculates the overall percent conservation in 100 bp surrounding base pair X. It repeats this for base pair Y, Z, and so on. The calculated percent conservations at base pairs X, Y, Z, etc. are then displayed and can be interpreted. Regions with higher than 70% conservation are considered highly conserved. Within the graph for each *rsp*, blue regions represent coding regions with high conservation. Red regions represent high conservation in regions previously annotated as non-coding that I found contain alternatively spliced exons that introduce PTCs. The gene map below the graph is the pattern of alternative splicing I observed for each *rsp*, and the red exons represent the alternatively spliced exons.

Figure 2.2

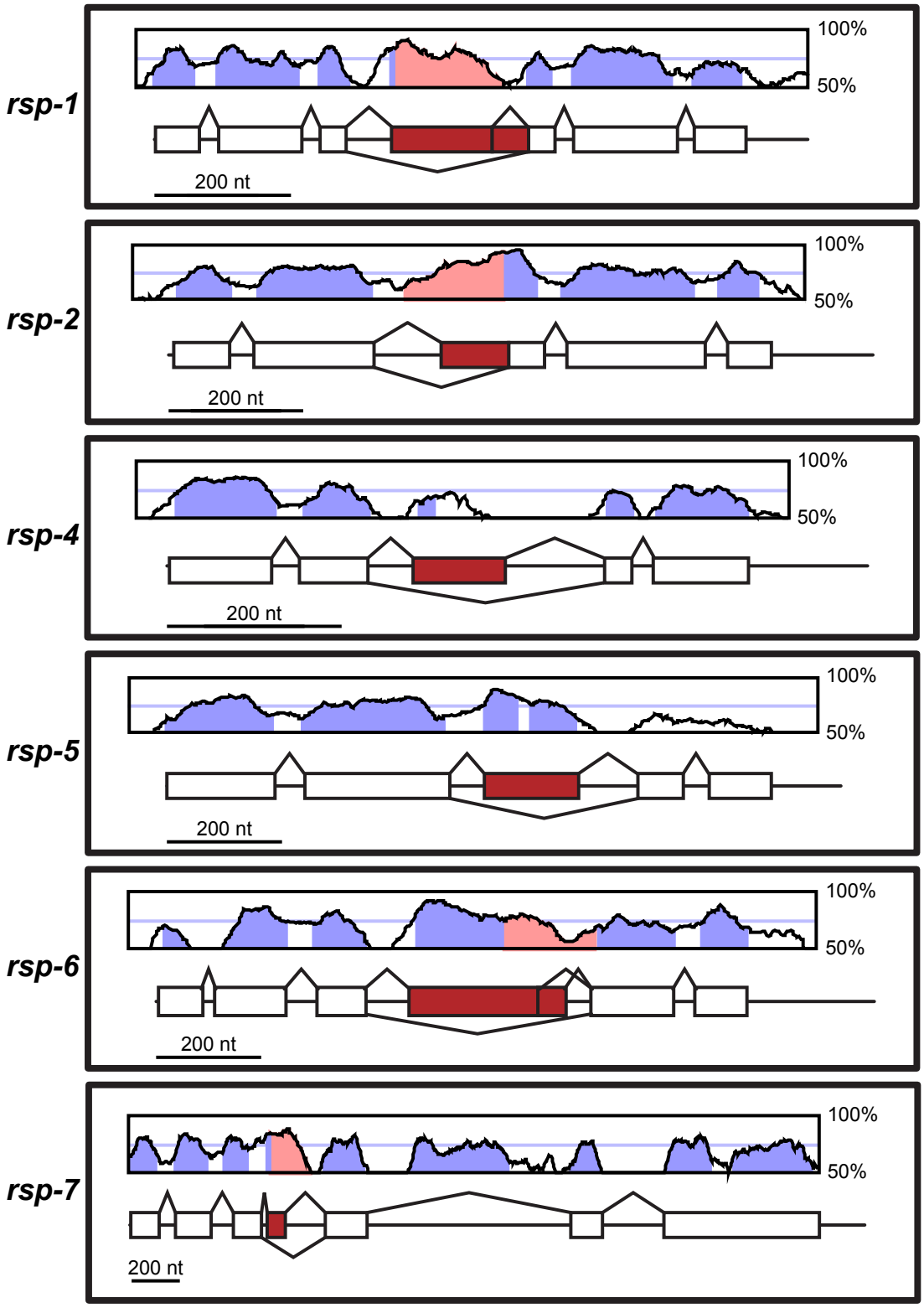

rsp-1
100%
50%
200 nt
rsp-2
100%
50%
200 nt
rsp-4
100%
50%
200 nt
rsp-5
100%
50%
200 nt
rsp-6
100%
50%
200 nt
rsp-7
100%
50%
200 nt

Figure 2.3. *rsp* qRT-PCR primers are designed to amplify specific isoforms.

qRT-PCR primers were designed to amplify *rsp(PTC), rsp(+),* or *rsp(total)* mRNAs. All primer pairs cross exon-exon junctions and generate only one product. Dashed parts of lines indicate regions that a primer spans across an exon-exon junction but does not anneal to. Solid lines represent regions that do anneal to cDNA and an arrowhead denotes the 5' to 3' direction of each primer. The primer design shown is for *rsp-5,* but is representative of how *rsp-1, rsp-2, rsp-4, rsp-6,* and *rsp-7* qRT-PCR primers are all designed.

Figure 2.3

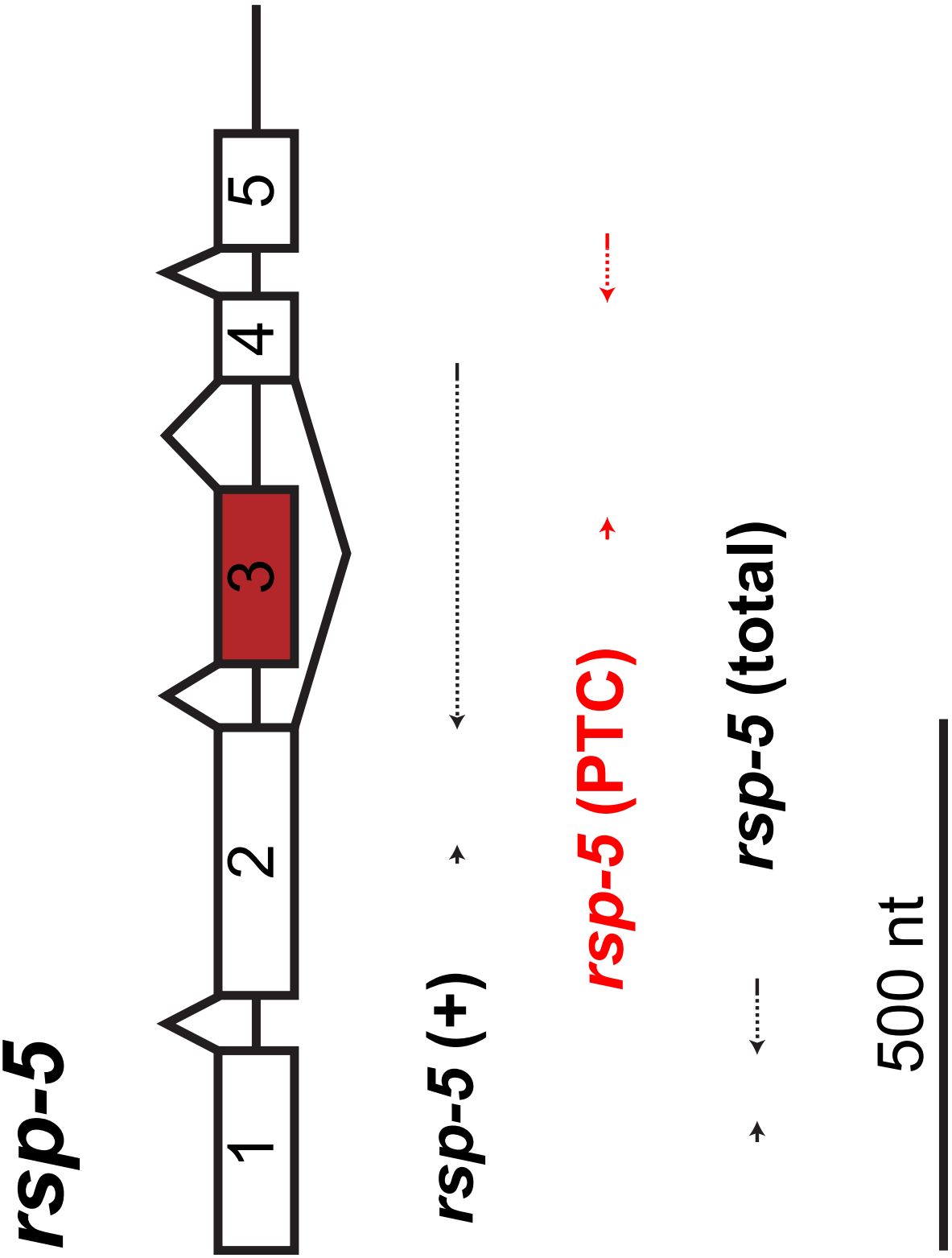


Figure 2.4. *rsp* mRNA expression is downregulated between embryo and L1.

qRT-PCR of developmentally staged A) wild-type samples show *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6*, and *rsp-7* (+) and (total) mRNAs decrease between embryo and L1 and remain low relative to embryo throughout development. B) *rsp(+)* and *rsp(total)* mRNA expression is similarly affected in *smg-2(-)* mutants. In *smg-2(-)* mutants, I also measured the mRNA expression for *rsp(PTC)* transcripts. *rsp(PTC)* transcripts overall show a similar, though less dramatic decrease in relative expression between embryo and L1. *Rsp(PTC), rsp(+),* and *rsp(total)* mRNAs were normalized to *eft-3*. Error bars indicate +/- S.D. n = 2-3 measurements.

Figure 2.4A. Regulation of *rsp* mRNAs in wild type.

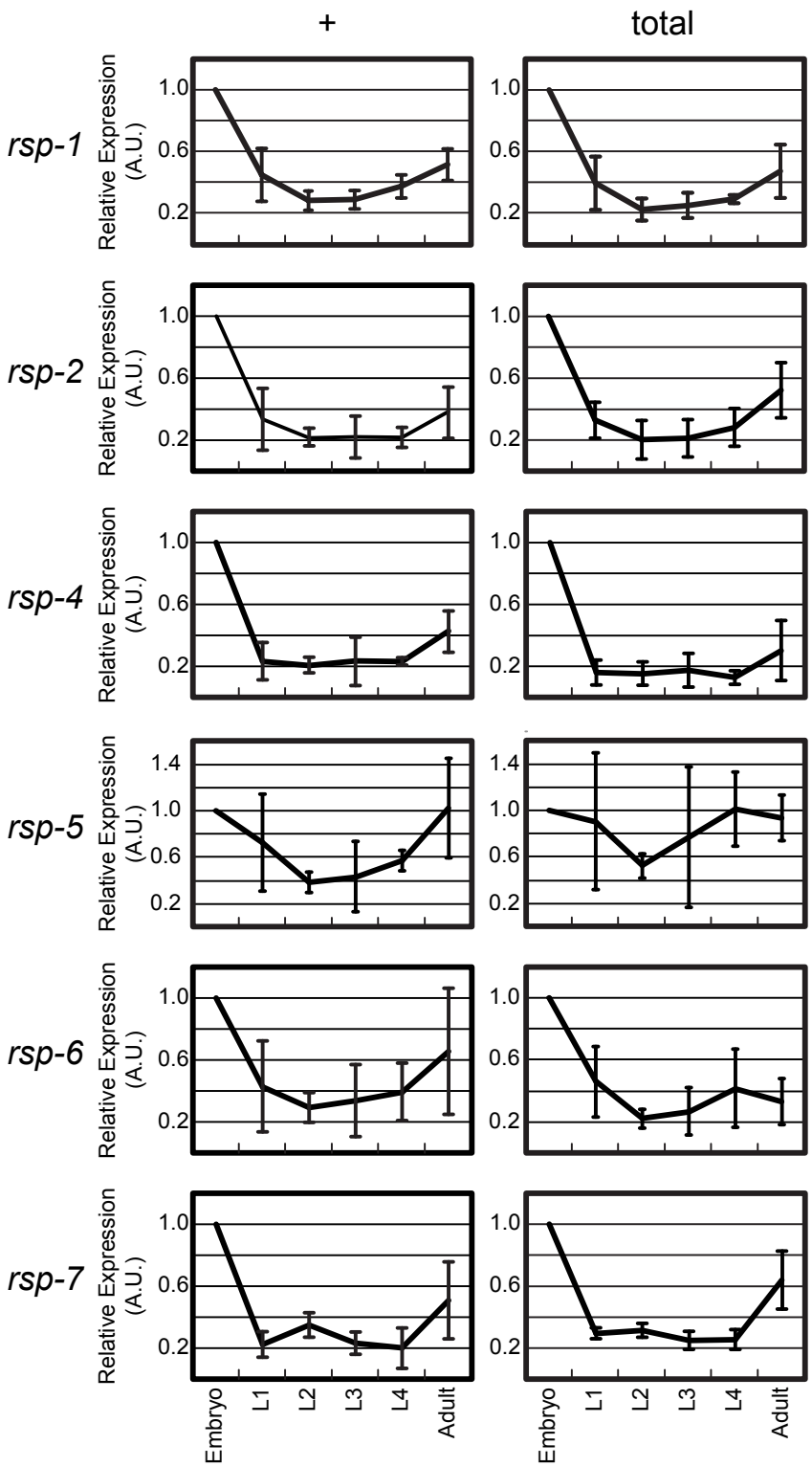


Figure 2.4B. Regulation of *rsp* mRNAs in *smg-2(-)* mutants.

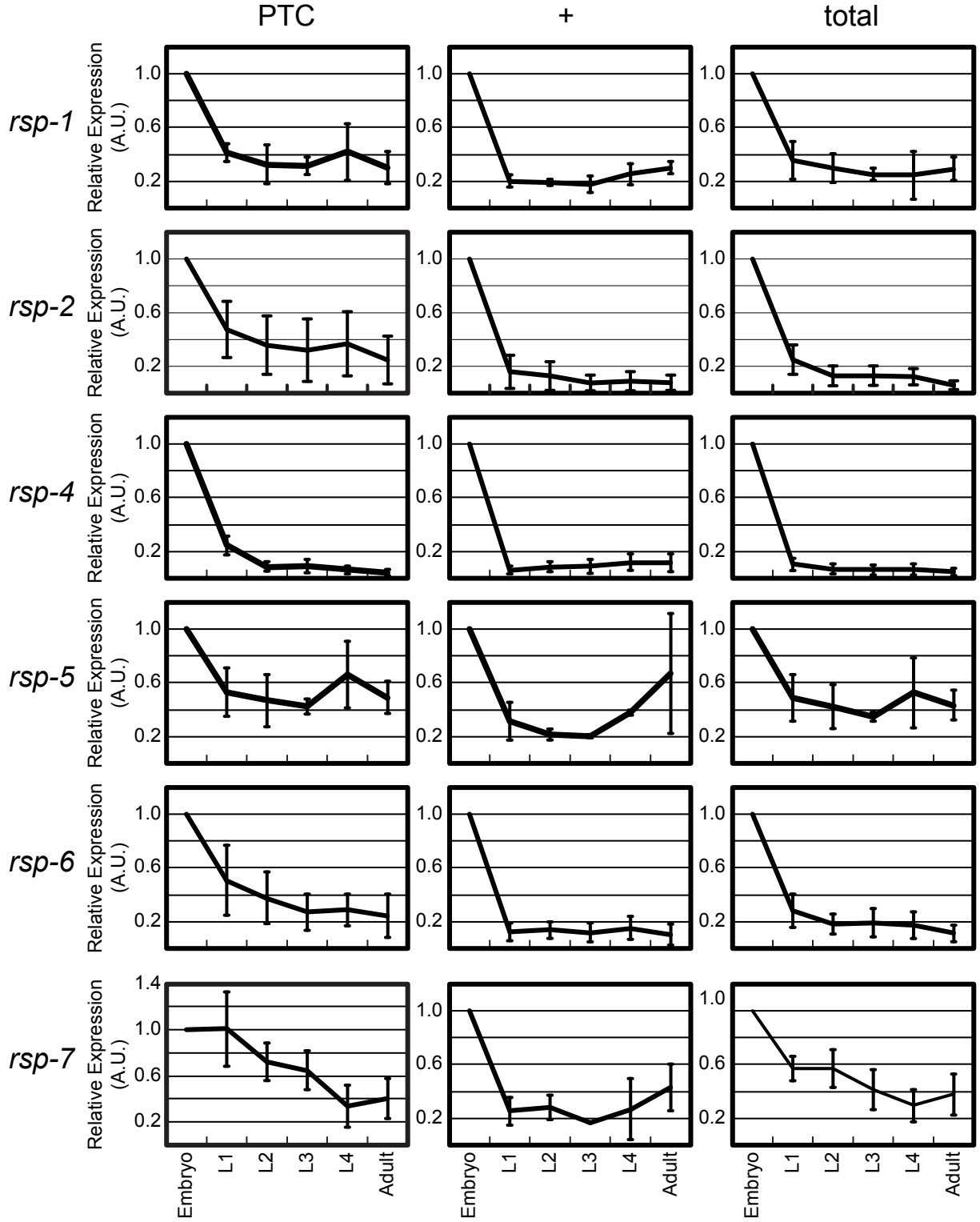


Figure 2.5. The proportion of *rsp(PTC)* mRNA changes throughout development in 4 of 6 alternatively spliced *rsp* mRNAs.

I calculated the proportion of *rsp(PTC)* mRNA [*rsp(PTC)/*sum of *rsp(PTC)* and *rsp(+)*] throughout development in *smg-2*(-) mutants for *rsp-1, -2, -4, -5, -6*, and -*7*. For *rsps-2, 4, -6*, and -*7* (note the scale changes in each chart), I observe a significant increase in the proportion of *rsp(PTC)* between the embryo to L1 transition. From L1 through adult, the proportion of *rsp-2(PTC)* remains consistent. The proportion of *rsp-4(PTC)* declines back to embryo levels in L2 then continues to decrease. The proportion of *rsp-6(PTC)* mRNA decreases from L1 to adult. The proportion of *rsp-7(PTC)* mRNA is consistent from L1 to L3 then decreases back to embryo levels. Error bars indicate +/- S.D. n=3.

Figure 2.5

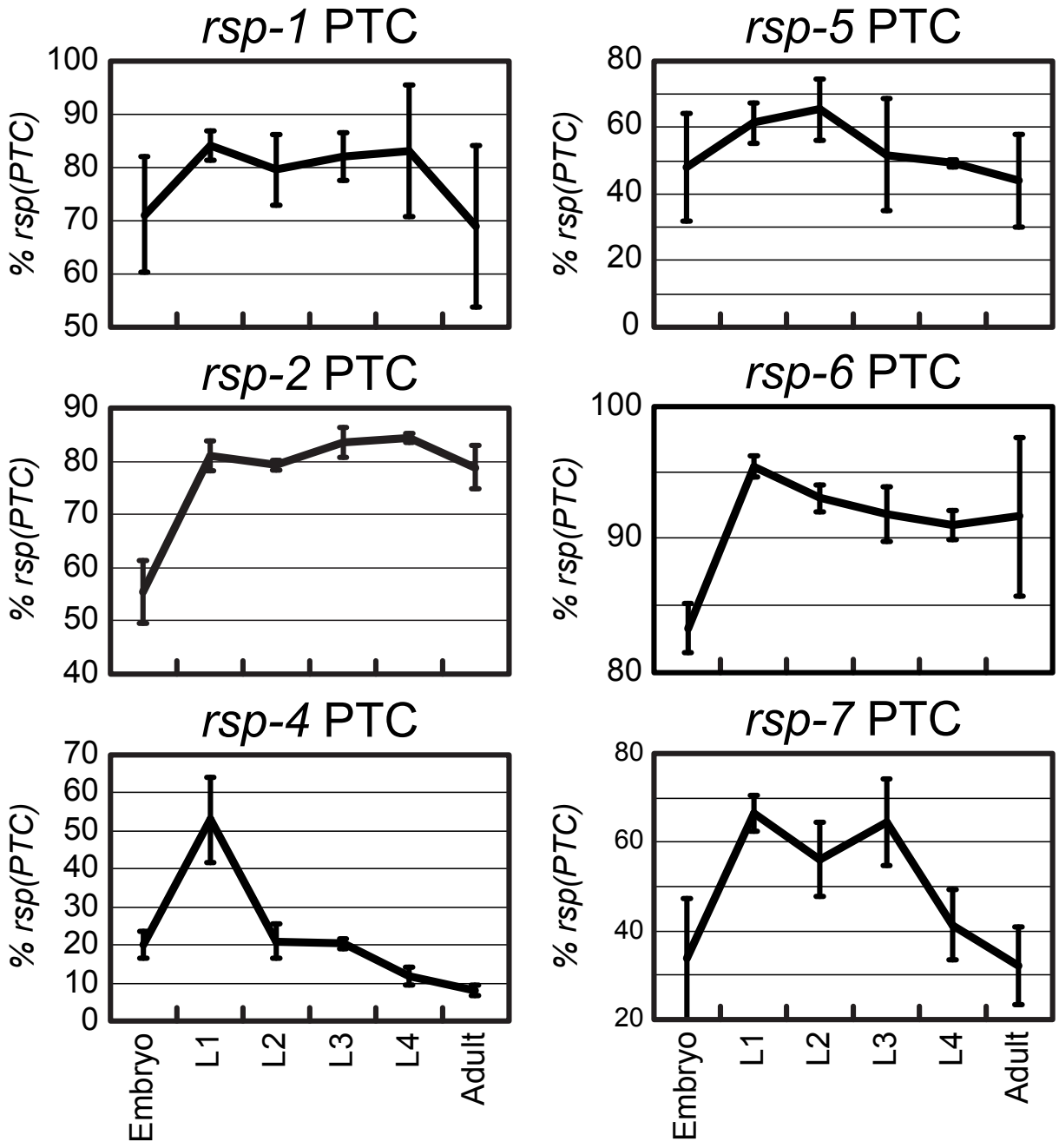

rsp-1 PTC
rsp-5 PTC
rsp-2 PTC
rsp-6 PTC
rsp-4 PTC
rsp-7 PTC
% rsp(PTC)
Embryo
L1
L2
L3
L4
Adult

Figure 2.6. Northern blots in wild type and *smg-2(-)* mutants show similar patterns as qRT-PCR.

A) Northern blots were performed on the first three developmental stages in both wild type (N2) and *smg-2(-)* mutants to measure *rsp-1, rsp-2, rsp-4, rsp-5,* and *rsp-6* mRNA expression normalized to *eft-3* expression. B) Quantifications of northern blots show changes in the relative expression of *rsp* mRNAs similar to those observed in qRT-PCR experiments. C) The proportion of *rsp(PTC)* mRNAs in the first three developmental stages measured by northern blots in *smg-2(-)* mutants is similar to results obtained with qRT-PCR. Error bars indicate +/- standard deviations, n=2.

Figure 2.6A

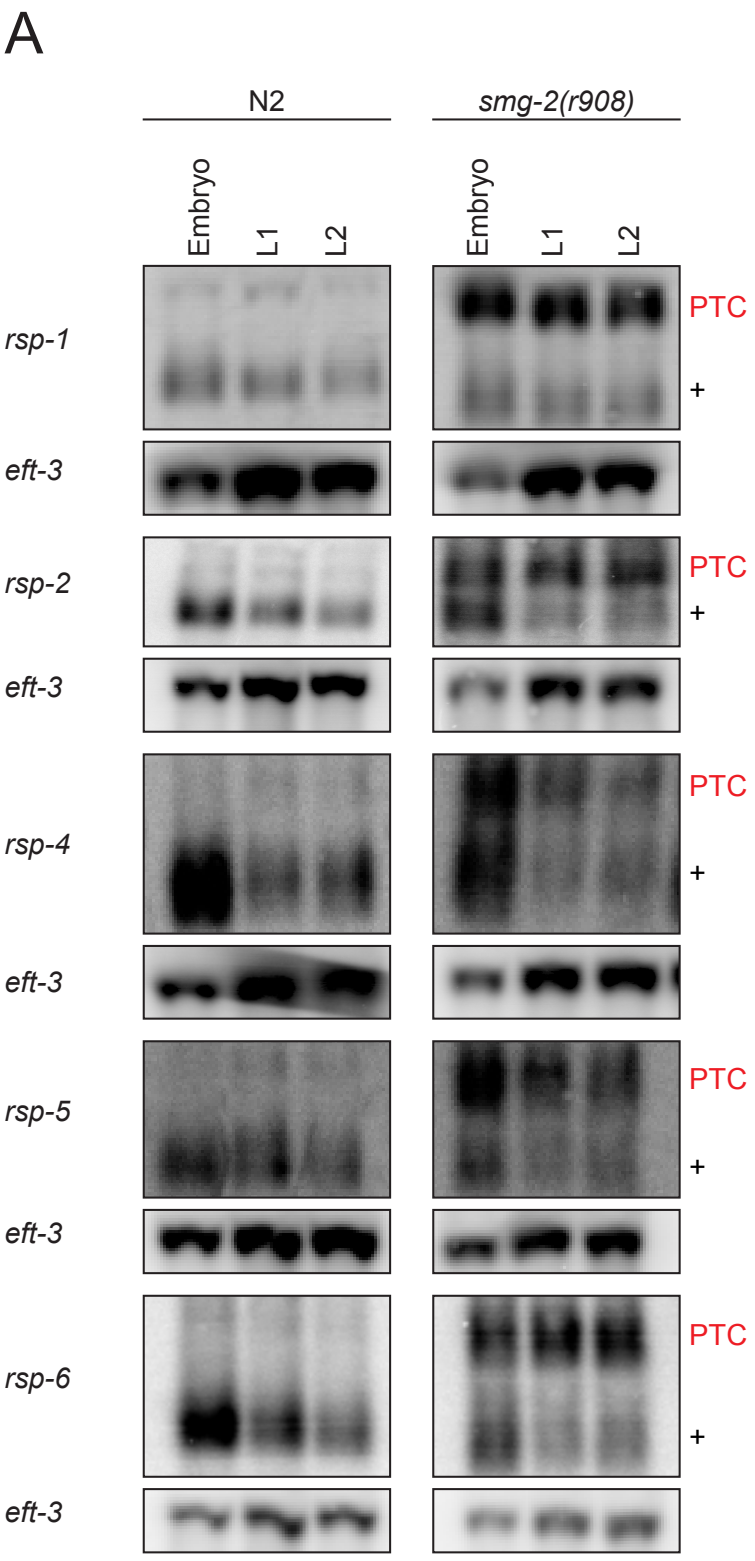

A
N2
smg-2(r908)
Embryo
L1
L2
Embryo
L1
L2
rsp-1
PTC
+
eft-3
rsp-2
PTC
+
eft-3
rsp-4
PTC
+
eft-3
rsp-5
PTC
+
eft-3
rsp-6
PTC
+
eft-3

Figure 2.6 Continued

B

N2

mRNA	Stage	+	total
rsp-1	Embryo	1	1
	L1	0.48 ± 0.11	0.49 ± 0.11
	L2	0.35 ± 0.01	0.38 ± 0.01
rsp-2	Embryo	1	1
	L1	0.30 ± 0.09	0.30 ± 0.08
	L2	0.20 ± 0.08	0.20 ± 0.07
rsp-4	Embryo	1	1
	L1	0.17 ± 0.03	0.17 ± 0.03
	L2	0.22 ± 0.03	0.22 ± 0.03
rsp-5	Embryo	1	1
	L1	0.47 ± 0.14	0.47 ± 0.14
	L2	0.46 ± 0.15	0.46 ± 0.15
rsp-6	Embryo	1	1
	L1	0.63 ± 0.52	0.58 ± 0.43
	L2	0.66 ± 0.66	0.57 ± 0.47

smg-2(-)

mRNA	Stage	PTC	+	total
rsp-1	Embryo	1	1	1
	L1	0.33 ± 0.07	0.27 ± 0.06	0.31 ± 0.07
	L2	0.35 ± 0.01	0.30 ± 0.01	0.34 ± 0.01
rsp-2	Embryo	1	1	1
	L1	0.43 ± 0.06	0.23 ± 0.12	0.34 ± 0.10
	L2	0.44 ± 0.03	0.22 ± 0.09	0.34 ± 0.04
rsp-4	Embryo	1	1	1
	L1	0.28 ± 0.03	0.13 ± 0.03	0.20 ± 0.02
	L2	0.20 ± 0.04	0.16 ± 0.05	0.18 ± 0.05
rsp-5	Embryo	1	1	1
	L1	0.47 ± 0.08	0.36 ± 0.02	0.43 ± 0.05
	L2	0.42 ± 0.06	0.41 ± 0.005	0.41 ± 0.04
rsp-6	Embryo	1	1	1
	L1	0.57 ± 0.02	0.22 ± 0.01	0.42 ± 0.04
	L2	0.60 ± 0.11	0.26 ± 0.01	0.45 ± 0.05

C

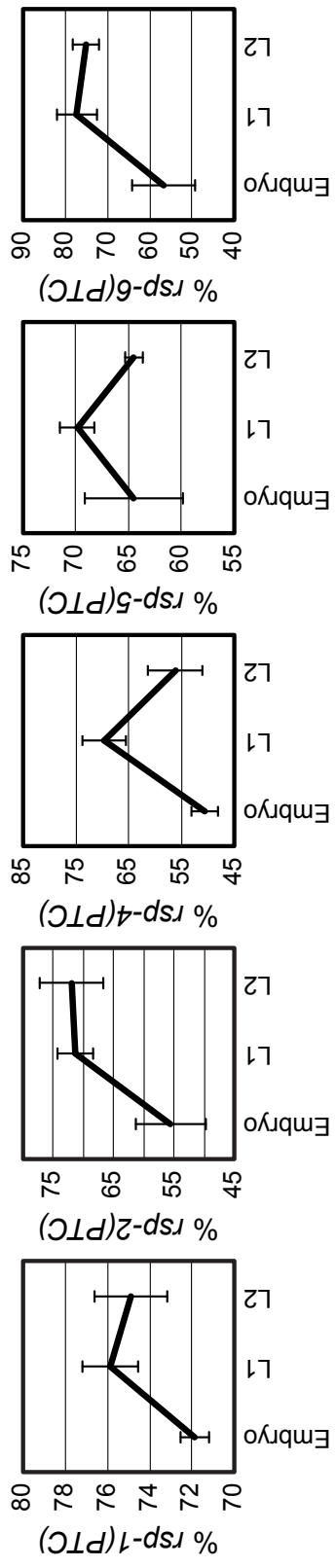


Figure 2.7. Semi-quantitative RT-PCR shows *rsp-7* splicing patterns change between embryo and L1.

Semi-quantitative RT-PCR shows a change in the splicing pattern of *rsp-7* during early development. My results with semi-quantitative RT-PCR are similar to qRT-PCR. Primers used to measure *rsp-7* splicing patterns are the same as those used in Figure 2.1. Error bars indicate +/- standard deviations, n=2.

Figure 2.7

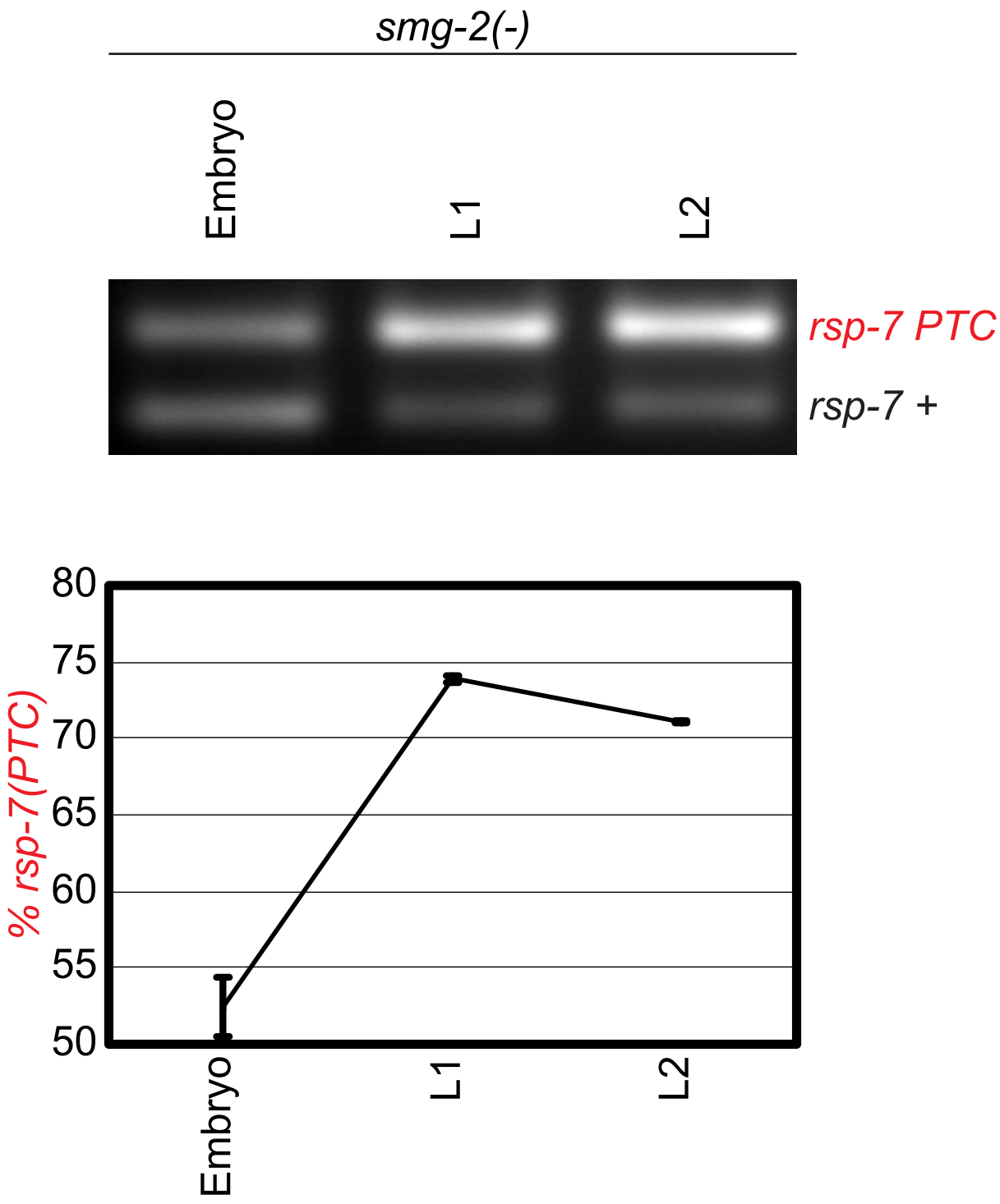

smg-2(-)
Embryo
L1
L2
rsp-7 PTC
rsp-7 +
80
75
70
65
60
55
50
% rsp-7(PTC)
Embryo
L1
L2

Figure 2.8. An anti-RSP-6 antibody specifically detects RSP-6 protein.

I generated and purified anti-RSP-6 polyclonal antibodies that detect an approximately 20 kDa protein absent in *rsp-6(ok798)* mutants. This western blots shows increased RSP-6 levels in increasing concentrations of mixed-stage wild-type (N2) lysate (N2 total protein lysate). Anti-actin was used as a loading control. Controls to show anti-RSP-6 specificity were collected by hand and represent 200 adult-stage hermaphrodites of either wild type (N2), *nT1/rsp-6(ok798)* heterozygotes (this strain carries the nT1 balancer chromosome described in Chapter 3), or *rsp-6(ok798)* homozygotes that do not express RSP-6 protein. OP50 is an *E. coli* food source for *C. elegans* and is presented as a negative control.

Figure 2.8

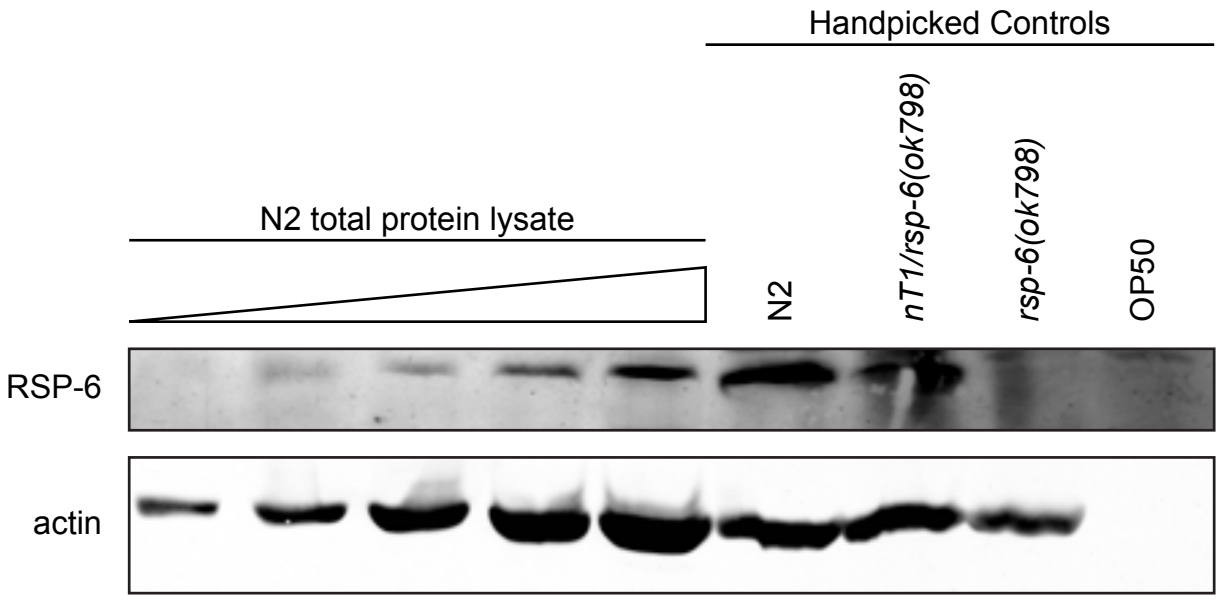

Handpicked Controls
N2 total protein lysate
N2
nT1/rsp-6(ok798)
rsp-6(ok798)
OP50
RSP-6
actin

Figure 2.9. RSP-6 is developmentally regulated in both wild type and *smg-2(-)* mutants.

RSP-6 expression in A) wild type (N2) B) *smg-2(-)*. RSP-6 is highly expressed during the first three developmental stages, then decreases dramatically between the L2 and L3 transition. RSP-6 returns to levels observed in embryo by adult stage. These expression patterns roughly parallel the *rsp-6* mRNA expression patterns. N2 blot is a representative of 3 blots; *smg-2*(-) blot is representative of 3-6 blots.

Figure 2.9

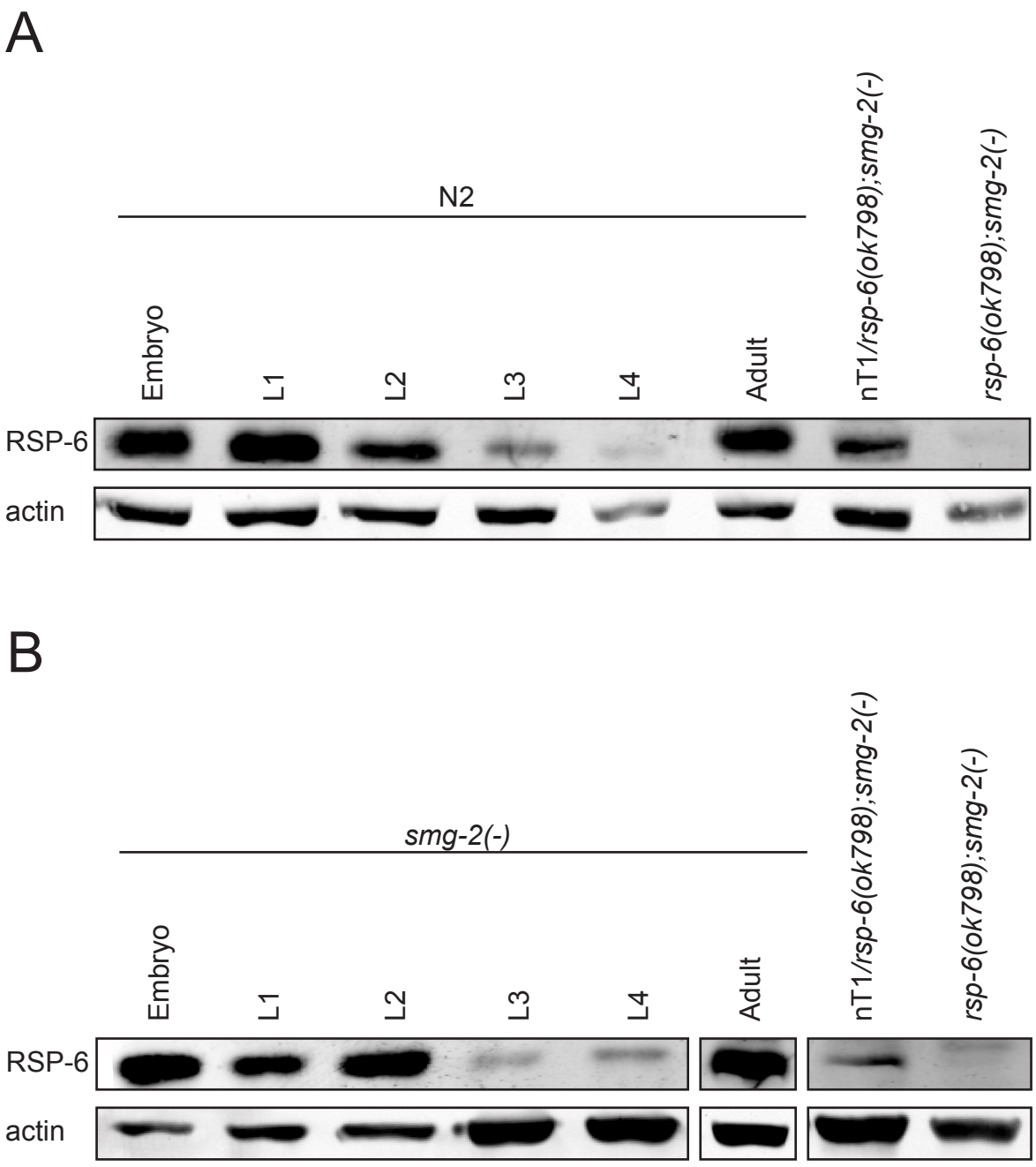

A
N2
Embryo
L1
L2
L3
L4
Adult
nT1/rsp-6(ok798);smg-2(-)
rsp-6(ok798);smg-2(-)
RSP-6
actin
B
smg-2(-)
Embryo
L1
L2
L3
L4
Adult
nT1/rsp-6(ok798);smg-2(-)
rsp-6(ok798);smg-2(-)
RSP-6
actin

Figure 2.10. Models of changes in *rsp* total mRNA and splicing patterns during the embryo to L1 transition.

A) Between embryo and L1, the rate of Pol II transcription could decrease dramatically, resulting in less mature total *rsp* mRNA. Changes in the rate of Pol II transcription could, but might not necessarily, also influence the ways in which *rsp* mRNAs are spliced. Thus, a decrease in Pol II transcription could correlate with an increase in the proportion of *rsp(PTC)* mRNA made through alternative splicing. B) Pol II transcription rates may not change between embryo and L1. The decrease in *rsp* total mRNA might therefore be explained by some other type of mRNA degradation. Both 2.10A and 2.10B depict changes in *rsp-6* mRNA expression and splicing patterns, but these models might also be applied to the regulation of *rsp-2, rsp-4,* and *rsp-7*.

Figure 2.10A

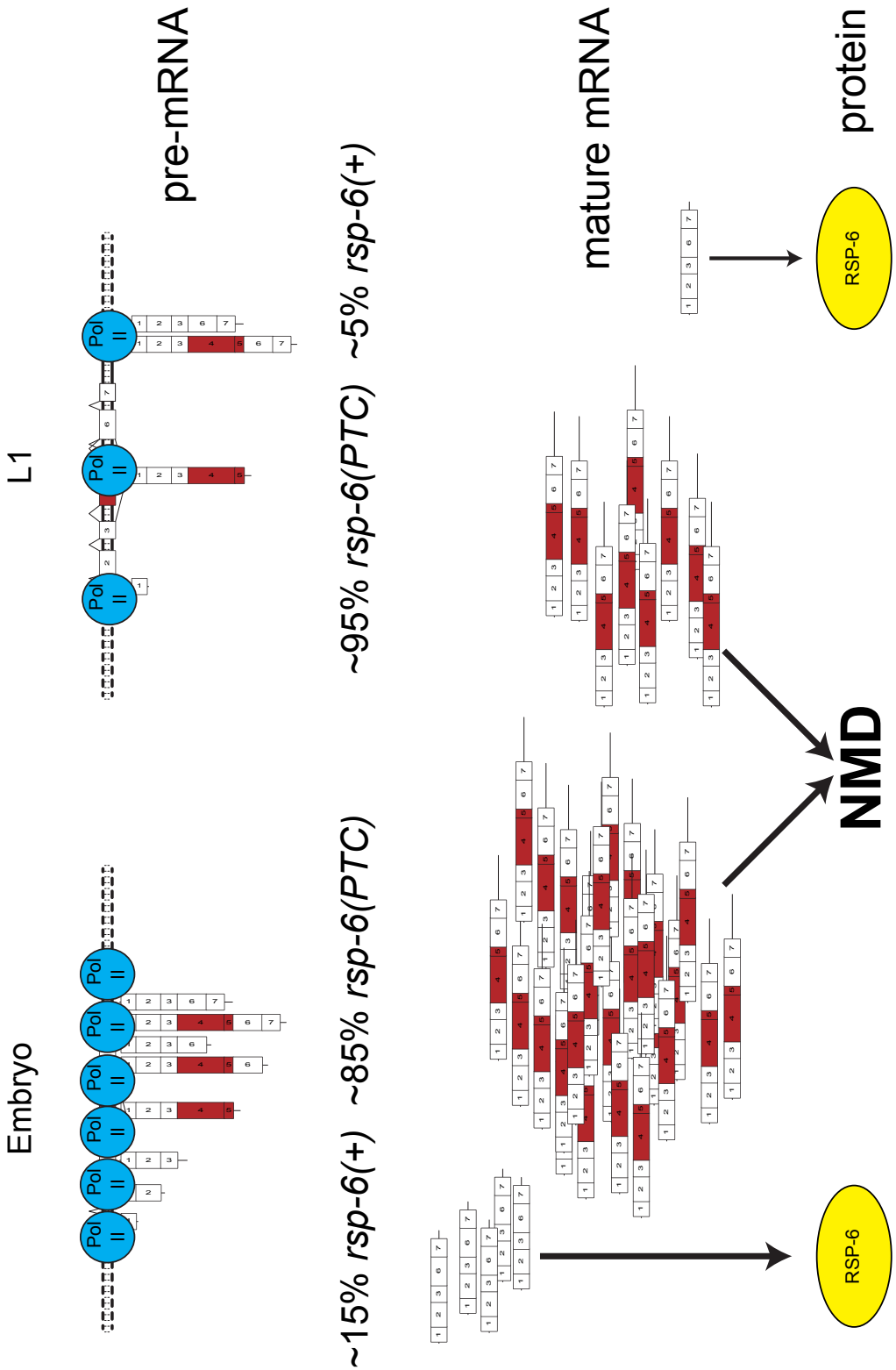
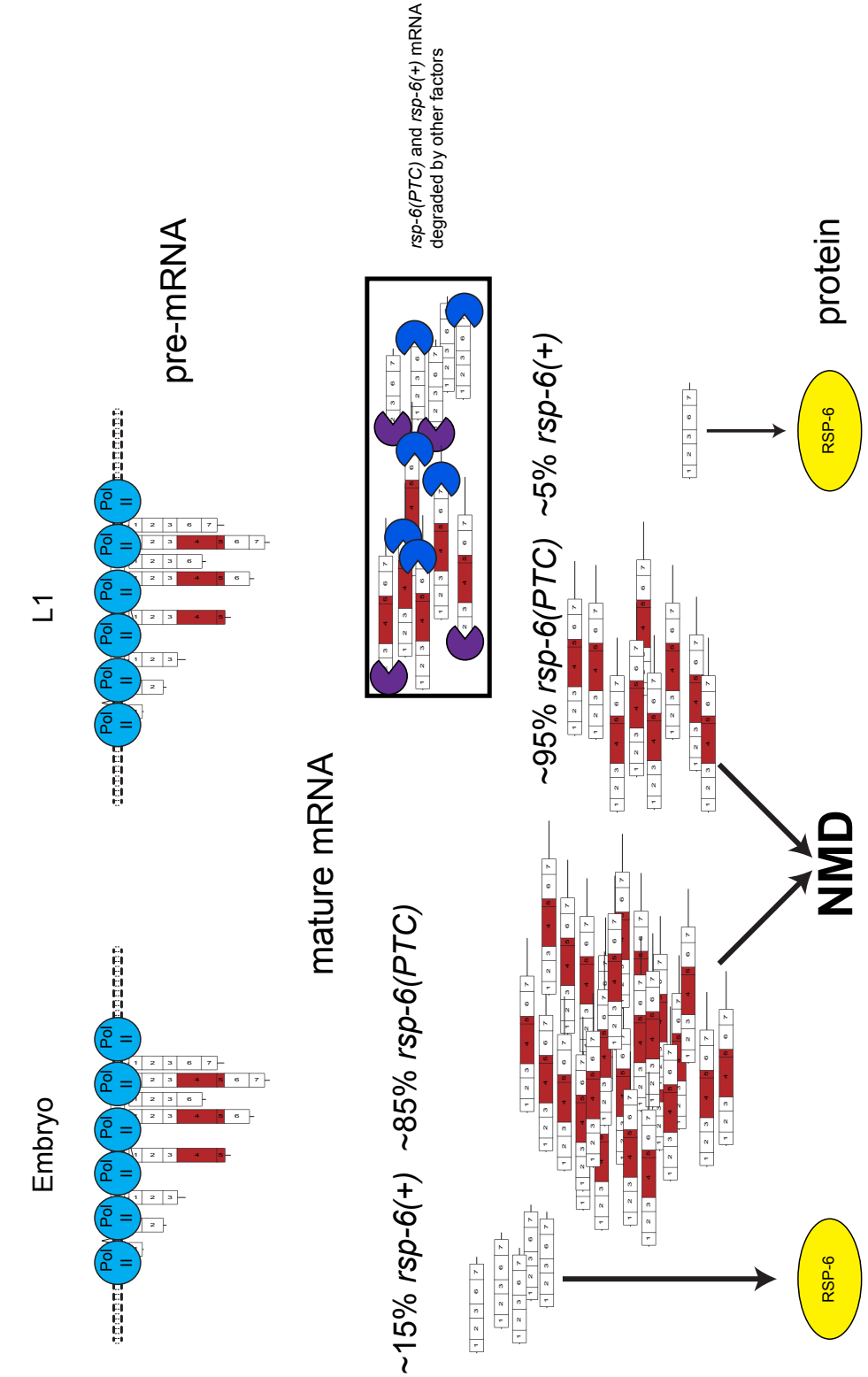

pre-mRNA
mature mRNA
protein
L1
Embryo
~95% rsp-6(PTC)
~5% rsp-6(+)
~15% rsp-6(+)
~85% rsp-6(PTC)
NMD
RSP-6
RSP-6

Figure 2.10B


Embryo
L1
pre-mRNA
mature mRNA
rsp-6(PTC) and rsp-6(+) mRNA degraded by other factors
~15% rsp-6(+) ~85% rsp-6(PTC)
~95% rsp-6(PTC) ~5% rsp-6(+)
NMD
RSP-6
RSP-6
protein

Chapter 3: Regulated Unproductive Splicing and Translation (RUST) of *C. elegans rsp* mRNAs

Abstract

Eukaryotic pre-mRNAs can be alternatively spliced to introduce premature termination codons (PTCs) in their open reading frames. Such mRNAs are substrates of the nonsense-mediated mRNA decay (NMD) pathway. Many of these alternative splicing events appear to be tightly regulated. Previous research in the Anderson and other labs described mRNAs whose splicing patterns are regulated by the abundance of the proteins encoded by those same mRNAs or proteins of related mRNAs (reviewed in [47]). A general model termed regulated unproductive splicing and translation (RUST) proposes that expression of alternatively spliced PTC-containing mRNAs is tightly regulated [157,158]. An extension of this model suggests other proteins can also influence the alternative splicing patterns of PTC-containing mRNAs. Though many examples of RUST exist, regulators of alternative splicing appear to be a prominent class of mRNAs subject to RUST. Such mRNAs include the serine/arginine-rich (SR) proteins and the heterogeneous nuclear ribonucleoprotein (hnRNP) families. Several hnRNP and SR proteins are known to regulate the alternative splicing of their own pre-mRNAs, described as autoregulation (reviewed in [47]). Some hnRNPs are known to regulate the alternative splicing of *hnRNP* paralogs through cross-regulation [104,162]. Cross-regulation has been defined as one or more proteins affecting the splicing patterns of a pre-mRNA that encodes a different protein.

Because *C. elegans rsp* mRNAs are alternatively spliced in patterns suggestive of RUST, I investigated whether *rsp* mRNAs are regulated by RUST. I generated strains in which expression of individual RSP proteins is either decreased or increased, and studied splicing patterns of all *rsp* genes possibly regulated by RUST. My results define two types of regulation: (i) Individual RSP proteins regulate the expression of *rsp(+)* vs. *rsp(PTC)* of their own splicing (autoregulation); and (ii) individual RSP proteins regulate splicing of other *rsp* genes (cross-regulation). My results identify a network of regulatory interactions among *rsp* genes and identify examples of redundancy and/or additivity of RSP function.

Introduction

Regulated Unproductive Splicing and Translation (RUST)

Mature mRNAs containing premature termination codons (PTCs) are common in many eukaryotes. Such mRNAs are rapidly degraded by the nonsense-mediated mRNA decay (NMD) pathway. NMD degrades PTC-containing mRNAs arising from a variety of processes, one of which is alternative splicing. The term "regulated unproductive splicing and translation" (RUST) describes a process in which mRNAs are intentionally spliced to include PTCs in order to reduce abundance of the corresponding mRNAs that are "full length" [157,158].

The broad significance of RUST is uncertain, and estimates of how many mRNAs are subject to regulation via RUST are varied. Early studies of EST databases suggested that 35% of alternatively spliced human genes are spliced such that one or more mRNAs contain PTCs [158]. Microarray profiling of selected alternative splicing events in human cell lines

suggested few instances of splicing to generate PTCs, but these experiments utilized a limited number of mRNAs and the selection of these mRNAs may have been biased [164]. High-throughput sequencing of *Arabidopsis* alternatively spliced mRNAs indicates that nearly 80% of alternatively spliced genes yield one or more mRNAs containing PTCs, that could be substrates for NMD [7].

Numerous examples of RUST have been described, and they involve two types of regulation: (i) events in which alternative splicing is regulated by the protein(s) encoded by that very gene (autoregulation), and (ii) events in which alternative splicing is regulated by proteins encoded by *other* genes in the genome (cross-regulation or networked regulation). In both autoregulation and cross-regulation, alternative splicing produces mature mRNAs that contain PTCs. I describe in this chapter examples of both autoregulatory and cross-regulatory RUST and my investigations of whether the splicing of *rsp* mRNAs is governed by RUST.

Autoregulation by RUST

The discovery of splicing to generate PTC-containing mRNAs of four different ribosomal proteins in the Anderson lab was one of the first clear examples of RUST [159]. *Rpl-12*, *rpl-3*, *rpl-7a*, and *rpl-10a encode C. elegans* large-subunit ribosomal proteins. All are spliced in patterns similar to the *rsp* mRNAs described above. Inclusion of specific exons in *rpl* mRNAs introduces PTCs to the mRNAs, causing them to be unstable in *smg(+)* strains. Such mRNAs are stable in *smg(-)* (NMD-defective) mutants. In strains that overexpress *rpl-12(+)* mRNA, the pattern of *rpl-12(+)* vs. *rpl-12(PTC)* splicing shifts, with an increased proportion of *rpl-12(PTC)* mRNA. Such observations support a model in which RPL-12, the

protein encoded by *rpl-12(+)* mRNA, either negatively regulates splicing to yield *rpl-12(+)* or positively regulates splicing to yield *rpl-12(PTC)*. Remarkably, generation of PTC-containing *rpl-3* and *rpl-12* mRNAs is conserved in mammalian species. Overexpression of rat *rpl-3(+)* mRNAs in rat cell lines shifts the *rpl-3* splicing pattern to increase the proportion of *rpl-3(PTC)* mRNA, similar to *rpl-12(+)* overexpression in *C. elegans* [160].

RUST influences splicing of many mRNAs, especially those encoding splicing factors (reviewed in [47]). mRNAs of two major classes of factors that regulate alternative splicing, the SR proteins and the hnRNP proteins, are alternatively spliced in a manner that implicates RUST in regulation of their expression. SRSF2 (SC35) is a canonical SR protein family member and its alternative splicing patterns are regulated by RUST. Alternative splicing of the *SRSF2* 3' UTR generates transcripts unstable in NMD-competent HeLa cells [161]. Unstable transcripts contain PTCs as a result of alternative splicing and are thus likely targets of NMD. Overexpression of SRSF2 increases the proportion of alternatively spliced *SRSF2* containing PTCs. *hnRNP L* encodes human hnRNP L protein. *hnRNP L* undergoes alternative splicing to include or exclude exon 6A [162]. Inclusion of exon 6A yields mature mRNA containing a PTC. RNAi knockdown of hnRNP L decreases the proportion of *hnRNP L* transcripts containing exon 6A, a finding consistent with a model of RUST. hnRNP L also binds its own pre-mRNA near exon 6A at CA clusters and therefore directs the splicing pattern of its own mRNA.

Cross-regulation by RUST

RUST is a model by which regulated alternative splicing coupled to rapid degradation of PTC-containing mRNAs contributes to protein homeostasis. This model can and has been extended to instances where a protein controls the regulated alternative splicing of a different pre-mRNA, contributing to the protein homeostasis of a different protein. Fewer examples of RUST via cross-regulation exist than via autoregulation, and those that have been described are limited to regulation of splicing regulators.

Human *hnRNP L* pre-mRNA is alternatively spliced to include or exclude an exon (exon 6A) that contains a PTC. *hnRNP L-like (hnRNP LL)* is similar to *hnRNP L* and undergoes similar alternative splicing to include or exclude an exon (exon 6A) that also contains a PTC. Splicing of *hnRNP LL* is regulated by a feedback loop similar to that of *hnRNP L*, but splicing of *hnRNP LL* can also be regulated by hnRNP L. Knockdown of hnRNP L protein by RNAi in HeLa cells decreases inclusion of the PTC-containing exon 6A in mature *hnRNP LL* mRNA. These observations suggest that hnRNP L acts to regulate inclusion or exclusion of exon 6A in the mature mRNAs of both *hnRNP L* and *hnRNP LL*. hnRNP L may do so directly by binding to CA clusters near exon 6A [162].

Human PTB (polypyrimidine tract binding) protein, also known as hnRNP I, is a repressive splicing regulator. PTB autoregulates splicing of its own mRNA to repress inclusion of exon 11, thereby yielding mature *PTB* mRNAs that contain PTCs [104]. A neuronal paralog of PTB, *nPTB*, is alternatively spliced at exon 10 and, like *PTB*, exclusion of exon 10 yields mature *nPTB* mRNAs that contain PTCs [104]. Knockdown of PTB in HeLa

cells results in decreased skipping of *nPTB* exon 10. Thus, PTB regulates the splicing pattern of *nPTB*.

Auto- and cross-regulation by RUST have been previously described for several mRNAs including those that are required for or regulate alternative splicing. Because alternatively spliced *C. elegans rsp* mRNAs often contain PTCs, I investigated whether alternative splicing of six different *rsp* mRNAs is autoregulated or cross-regulated (or both) by individual RSP proteins. This chapter describes my experiments to address these questions. I will describe a general approach designed to either reduce or increase abundance of individual RSP proteins, followed by my investigations of the effects such manipulations have on expression of each *rsp* gene.

Results

I constructed strains designed to increase or decrease expression of individual RSP proteins. To decrease RSP expression, I utilized *rsp* deletions [indicated as *rsp(Δ)*] obtained from either the Oklahoma or Japanese knockout consortium. Such deletion homozygotes are almost certainly complete null alleles (see below). Deletion heterozygotes are expected to reduce expression of the corresponding RSP protein, although the magnitude of the reduction is uncertain. I investigated splicing of *rsp* mRNAs in strains that are wild-type for all *rsp* genes [*rsp(+)*], heterozygous for an *rsp* deletion [*rsp(Δ)*/+], and homozygous for an *rsp* deletion [*rsp(Δ)*]. To increase RSP expression, I constructed strains in which an *rsp* transgene that cannot be spliced to produce *rsp(PTC)* mRNAs is expressed from a constitutively strong

promoter. I then investigated splicing of each *rsp* gene in overexpression transgenic strains and controls. I discuss the genetic manipulations in more detail below.

Rsp **deletion alleles**

I obtained deletion alleles of *rsp-2*, *rsp-4*, *rsp-5*, and *rsp-6* to reduce or eliminate RSP protein expression. Figure 3.1 diagrams *rsp* gene structures and regions deleted by *rsp-2(tm952), rsp-4(tm837), rsp-5(ok324),* and *rsp-6(ok798)*, abbreviated *rsp-2(Δ)*, *rsp-4(Δ)*, *rsp-5(Δ)*, and *rsp-6(Δ)* hereafter. Both *rsp-2(Δ)* and *rsp-4(Δ)* delete regions that include the AUG translation initiation codon. Both *rsp-5(Δ)* and *rsp-6(Δ)* delete nearly the entirety of their respective gene. I designed multiplexed genotyping primers that uniquely detect wild-type and deletion alleles in deletion heterozygotes or homozygotes (See Figure 3.1). Guided by these genotyping primers, I constructed a variety of deletion heterozygotes, deletion homozygotes, and deletion double mutants (see below). All *rsp* deletions were outcrossed to wild type at least six times prior to their analysis.

Heterozygous deletions were propagated as balanced heterozygotes using GFP-marked balancer chromosomes. *nT1[qIs51]* contains a GFP marker, is *rsp-6(+)*, and balances the *rsp-6* region. Similarly, *mIn1[mIs14]* contains a GFP marker, is *rsp-2(+)*, *rsp-4(+)*, and *rsp-5(+)* and balances the *rsp-2*, *rsp-4*, and *rsp-5* regions. *nT1* and *mIn1* are described in [177]. *nT1[qIs51]* homozygotes express GFP and die during larval development; *mIn1[mIs14]* homozygotes express pharyngeal GFP and have a Dpy phenotype. Among the self-fertilized progeny of *rsp-6(Δ)* / *nT1[qIs51]*, *rsp-6(Δ)* homozygotes are non-green and non-lethal, *rsp-6(Δ)* deletion heterozygotes are green and non-lethal, and *nT1* homozygotes

are lethal. Among the self-fertilized progeny of *rsp-2(Δ) / mIn1[mIs14]*, *rsp-4(Δ) / mIn1[mIs14]*, and *rsp-5(Δ) / mIn1[mIs14]*, *rsp(Δ)* deletion homozygotes are non-green and non-Dpy, *rsp(Δ)* deletion heterozygotes are green and non-Dpy, and *mnIn1* homozygotes are green and Dpy. *Rsp-6(Δ)* deletion homozygotes are sterile. *Rsp-2(Δ)* deletion homozygotes, *rsp-4(Δ)* deletion homozygotes, and *rsp-5(Δ)* deletion homozygotes have no visible phenotype.

I constructed *rsp(Δ)*; *rsp(Δ)* double mutants for all combinations of *rsp-2(Δ), rsp-4(Δ),* and *rsp-5(Δ)* alleles. *Rsp-2(Δ) rsp-4(Δ)* doubles, *rsp-2(Δ) rsp-5(Δ)* doubles, and *rsp-4(Δ) rsp-5(Δ)* doubles are all viable and fertile as homozygotes. To build double mutants with the *rsp-6(Δ)* allele, I constructed strains heterozygous for *rsp-6(Δ)* and homozygous for *rsp-2(Δ), rsp-4(Δ),* or *rsp-5(Δ)*. Strains heterozygous for *rsp-6(Δ)* are marked with the *nT1* balancer chromosome and allow the strain to be propagated since *rsp-6(Δ)* homozygotes are sterile. *nT1[qIs51]/rsp-6(Δ);rsp-4(Δ)* and *nT1[qIs51]/rsp-6(Δ);rsp-5(Δ)* strains were generated. These strains segregate heterozygous for *rsp-6(Δ)* and homozygous for either *rsp-4(Δ)* or *rsp-5(Δ),* or homozygous for *rsp-6(Δ);rsp-4(Δ)* and *rsp-6(Δ);rsp-5(Δ)*. Strains heterozygous for *rsp-6(Δ)* appear green. Strains homozygous for *rsp-6(Δ)* are non-green. Therefore, double *rsp-6(Δ);rsp-4(Δ),* or *rsp-6(Δ);rsp-5(Δ)* can be selected based on the absence of GFP expression. Strains homozygous for *rsp-6(Δ)* and either *rsp-4(Δ)* or *rsp-5(Δ)* were selected in the experiments that follow. Though viable, *rsp-6(Δ);rsp-4(Δ)* and *rsp-6(Δ);rsp-5(Δ)* are sterile. Finally, I also generated *nT1/rsp-6(Δ);rsp-2(Δ)* strains, however, I was never able to recover *rsp-6(Δ)* homozygotes from this strain. I conclude the combination of *rsp-6(Δ);rsp-2(Δ)* is likely a synthetic lethal. Remaining *rsp(Δ)* and combinations of *rsp(Δ)* mutants have

no visible phenotype. However, brood sizes of *rsp-2(Δ) rsp-5(Δ)* and *rsp-4(Δ) rsp-5(Δ)* double mutants are slightly reduced compared to wild type.

Transgenic strains that overexpress *rsp* genes

To increase expression of individual RSP proteins, I constructed strains in which *gfp::rsp* transgenes are expressed from a constitutively strong promoter. Such strains have the following properties (see Figure 3.2A): (i) Translation of the transgene mRNA fuses the COOH-terminus of GFP in frame to a full length RSP*(+)* protein; (ii) regions near *rsp* alternatively spliced exons were replaced with cDNA sequences, such that only *gfp::rsp(+)* mRNAs can be expressed; (iii) the 3' UTRs of *gfp::rsp* transgenes derive from *unc-54*, allowing endogenous *rsp* mRNAs to be distinguished by both size and sequence from transgene mRNAs; and (iv) expression of *gfp::rsp* transgenes is driven by the strong and constitutively active *eft-3* promoter. I established high-copy, extra chromosomal, transgene arrays by co-injection with *rol-6(su1006)*, a dominant marker that confers an easily distinguishable Roller phenotype.

Because neither *rsp-2(Δ)* nor *rsp-4(Δ)* single mutants have a visible phenotype, I could not test whether the transgene GFP::RSP proteins are functional by phenotypic rescue of *rsp-2* or *rsp-4* mutant phenotypes. However, both *gfp::rsp-2* and *gfp::rsp-4* transgenic strains express abundant exogenous mRNA (see below) and GFP::RSP-2 or GFP::RSP-4 fusion proteins of the expected molecular weights (see Figure 3.2B, lanes 3-5). Furthermore, two independent lines of *gfp::rsp-4* affect the splicing pattern of *rsp-4* in a manner opposite that of *rsp-4(Δ)* heterozygotes (see below). The *gfp::rsp-2* transgenic strain does not affect splicing

of *rsp-2*, but both *gfp::rsp-2* and *gfp::rsp-4* affect splicing of other *rsp* mRNAs in a manner opposite to that of *rsp-2(Δ)* and *rsp-4(Δ)* mutants (see below). I therefore conclude *gfp::rsp-2* and *gfp::rsp-4* transgenic strains likely produce at least partially functioning proteins.

***Gfp::rsp-6* transgenes may be silenced**

Two independent *gfp::rsp-6* transgenic strains express transgenic mRNAs and GFP::RSP-6 fusion proteins (Figure 3.2B, lanes 6-7). The abundance of GFP::RSP-6 fusion proteins, however is considerably lower than that of GFP, GFP::RSP-2, or GFP::RSP-4 transgenic strains expressed from the same promoter (Figure 3.2B, compare lanes 6-7 with lane 2-5). I tested whether my two *gfp::rsp-6* transgenes rescue the sterility phenotype of *rsp-6(Δ)* homozygotes. I constructed *rsp-6(Δ)/nT1[qIs51] gfp::rsp-6[rEx187]* and *rsp-6(Δ)/nT1[qIs51] gfp::rsp-6[rEx190]* and examined their self-fertilized progeny. If either *rEx187* or *rEx190* rescues the sterility of *rsp-6(Δ)* deletion homozygotes, offspring should include a class of non-green, array-containing, fertile animals (genotype *rsp-6(Δ); gfp::rsp-6[rEx187* or *rEx190]*). Non-green, array-containing animals were rarely observed, and I was only able to find approximately 15 offspring of that phenotype (15 total from both *rEx187* and *rEx190*), but none were of the expected genotype. Instead, all were of genotype *rsp-6(Δ)/nT1[qIs51]; rsp-6[rEx187* or *rEx190]*. The reasons for the failure of these strains to express GFP is unknown, but one possibility is that the presence of *rEx187* or *rEx190* elicited transgene silencing not only of GFP encoded on the extrachromosomal array but also GFP of *nT1[qIs51]*. However, the proportion of non-green *rsp-6(Δ)/nT1[qIs51]; rsp-6[rEx187* or *rEx190]* is quite small, therefore silencing may only impact a small fraction of *gfp::rsp-6*

transgenic worms. I conclude that neither *gfp::rsp-6[rEx187]* nor *gfp::rsp-6[rEx190]* express sufficient functional GFP::RSP-6 protein to rescue the sterility of *rsp-6(Δ)* homozygotes. Sterility, however, is inherently a germline defect, and obtaining expression of *C. elegans* transgenes in the germline is notoriously difficult [178].

Rsp splicing patterns and RUST

I describe regulation of *rsp* gene expression by alternative splicing in following sections. I crossed deletions and transgenes described above into *smg-2(-)* genetic backgrounds and quantified both *rsp(+)* and *rsp(PTC)* mRNAs in deletion heterozygotes, deletion homozygotes, and RSP overexpression strains. I first describe feedback autoregulation that governs expression of *rsp-2, rsp-4, rsp-5,* and *rsp-6* splicing. I then describe regulated alternative splicing that is affected by the absence of or overexpression of specific RSP proteins.

Autoregulation of *rsp* splicing patterns

Many alternatively spliced mRNAs are regulated by RUST to maintain protein homeostasis. I hypothesized that *rsp* splicing might be similarly regulated. I envision that individual RSP proteins directly or indirectly regulate alternative splicing of their own pre-mRNA. Under conditions where a surplus of an RSP protein is present, a greater proportion of *rsp* pre-mRNA would be spliced to include a PTC, thereby maintaining *rsp(+)* mRNA at appropriate levels. Conversely, under conditions where a deficit of an RSP protein is present, a greater proportion of *rsp* pre-mRNA would be spliced to NOT include a PTC, thereby

increasing *rsp(+)* mRNA at appropriate levels. I tested this model by investigating *rsp* deletion heterozygotes and *gfp::rsp* overexpression strains as described above.

Rsp deletion heterozygotes are expected to reduce expression of the corresponding RSP protein. Such gene-dose-dependent expression is common in eukaryotes, although the magnitude of the reduction is *a priori* difficult to predict. If expression of a specific RSP is reduced in an *rsp(Δ)* heterozygote, and if that RSP regulates alternative splicing of itself, then the deletion heterozygote will have altered patterns of *rsp* splicing. I tested this hypothesis by quantifying *rsp(+)* and *rsp(PTC)* mRNAs for *rsp-2, rsp-4, rsp-5,* and *rsp-6* in strains that are heterozygous for *rsp-2(Δ)*, *rsp-4(Δ)*, *rsp-5(Δ)*, or *rsp-6(Δ)*, respectively. Balancer chromosomes and phenotypes that allowed me to uniquely identify deletion heterozygotes are described above. I collected approximately 300, handpicked, L4-stage, deletion heterozygotes for each tested gene plus control samples of balancer heterozygotes that do NOT contain an *rsp* deletion [genotypes *nT1[qIs51]*/+ and *mIn1[mIs14]*/+]. All tested strains are additionally *smg-2(-)*. I quantified *rsp(+)*, *rsp(PTC)*, and *rsp(total)* mRNAs of *rsp-2*, *-4*, *-5*, and *-6* relative to *eft-3* mRNA in each strain. The results are shown in Figure 3.3 (A, C, E, and F) and discussed in more detail below.

Similarly, I tested whether demonstrated or putative overexpression of RSP-2, RSP-4, and RSP-6 (see above) affect *rsp* splicing. I quantified *rsp(total)* mRNA (includes both endogenous and exogenous sources) and endogenous (as opposed to transgene) *rsp(+)* and *rsp(PTC)* mRNAs of *rsp-2*, *-4*, and *-6* relative to *eft-3* mRNA. qRT-PCR primer pairs are specific for each endogenous *rsp(+)* mRNA and amplify from the *rsp* 3' UTR (see Figure 3.2A). Strains overexpressing RSP-2, RSP-4, and RSP-6 do not express *rsp(PTC)* mRNA,

thus *rsp(PTC)* mRNA in these strains is endogenous. I quantified mRNAs in handpicked L4-stage, *smg-2(-)* strains that contain *gfp::rsp-2[rEx188], gfp::rsp-4[rEx186* or *rEx191], gfp::rsp-6[rEx187* or *rEx190]*, or no transgene. My results are shown in Figure 3.3 (B, D, and G) and are discussed in more detail below.

Rsp-2 alternative splicing in *rsp-2* mutants

Splicing of *rsp-2* mRNA is altered in *rsp-2(Δ)/+* deletion heterozygotes compared to *rsp-2(+)* strains (Figure 3.3A). The proportion of *rsp-2(PTC)* mRNA decreases significantly ($p<0.05$) in deletion heterozygotes. I conclude that the proportion of *rsp-2(PTC)* in *rsp-2(Δ)/+* heterozygotes is lower than in *rsp-2(+)* strains.

Mutants that overexpress *gfp::rsp-2* mRNA show no statistically significant changes in *rsp-2* expression or alternative splicing (Figure 3.3B). Expression of *rsp-2(+)* in my only *gfp::rsp-2* transgenic strain is unaltered. Expression of *rsp-2(PTC)* and *rsp-2(total)* mRNAs appears to increase, but large error bars make the apparent increases not statistically significant.

Rsp-4 alternative splicing in *rsp-4* mutants

Splicing of *rsp-4* mRNA is altered in *rsp-4(Δ)/+* deletion heterozygotes compared to *rsp-4(+)* strains (Figure 3.3C). The relative expression of *rsp-4(total)* and *rsp-4(PTC)* mRNAs decrease significantly ($p<0.01$ and $p<0.05$, respectively), while the relative expression of *rsp-4(+)* mRNA remains unchanged or is perhaps slightly elevated. The proportion of *rsp-4(PTC)* mRNA decreases several fold ($p<0.05$). Such results indicate *rsp-4*

splicing is affected by the level of *rsp-4(total)* mRNA and are consistent with a model of RUST. Reducing the quantity of total mRNA alters splicing such that the relative expression of *rsp-4(+)* remains relatively unchanged.

Mutants that overexpress *gfp::rsp-4* mRNA affect *rsp-4* splicing in a manner that is opposite to that of *rsp-4(Δ)/+* heterozygotes (Figure 3.3D). The relative expression of *rsp-4(total)* mRNA increases 3- to 4-fold ($p<0.01$) in two independent *gfp::rsp-4* overexpression lines (note this measurement includes both endogenous and exogenous sources of *rsp-4(total)* mRNA, but all other measurements are of only endogenous *rsp-4* mRNA). The relative expression of *rsp-4(PTC)* mRNA increases approximately 1.6-fold ($p<0.05$ and $p<0.01$), while the relative expression of *rsp-4(+)* mRNA remains unchanged. The proportion of *rsp-4(PTC)* mRNA may increase, although the observed increases are not statistically significant. I conclude that elevated expression of *gfp::rsp-4* mRNA affects splicing of *rsp-4* mRNAs in a manner consistent with RUST. The decreased expression of *rsp-4(total)* mRNA in *rsp-4(Δ)/+* heterozygotes and the increased expression of *rsp-4(total)* mRNA in *gfp::rsp-4* overexpression strains suggest that changes in *rsp-4(total)* mRNA are responsible for changes in the balance of *rsp-4(PTC)* vs. *rsp-4(+)* splicing.

<u>*Rsp-5* alternative splicing in *rsp-5* mutants</u>

Splicing of *rsp-5* mRNA is altered in *rsp-5(Δ)/+* deletion heterozygotes compared to *rsp-5(+)* strains (Figure 3.3E). The relative expression of *rsp-5(total)* and *rsp-5(PTC)* mRNAs decrease approximately two-fold ($p<0.05$ for both), while the relative expression of *rsp-5(+)* mRNA remains unchanged. The proportion of *rsp-5(PTC)* mRNA decreases about

2-fold ($p<0.01$). As with *rsp-4*, such results suggest that *rsp-5* alternative splicing is affected by the relative expression of *rsp-5(total)* mRNA in a manner consistent with a model of RUST. *Gfp::rsp-5* fusions were difficult to clone, therefore I did not generate GFP::RSP-5 overexpression strains.

Rsp-6 alternative splicing in *rsp-6* mutants

Splicing of *rsp-6* mRNA is altered in *rsp-6(Δ)/+* deletion heterozygotes compared to *rsp-6(+)* strains (Figure 3.3F). The relative expression of *rsp-6(total)* and *rsp-6(PTC)* mRNAs decrease approximately 4-fold and 5-fold, respectively ($p<0.01$ for both), while the relative expression of *rsp-6(+)* mRNA is unchanged or perhaps slightly decreased. The proportion of *rsp-6(PTC)* mRNA decreases significantly ($p<0.01$). I conclude that splicing of *rsp-6* is affected by the relative expression of *rsp-6(total)* mRNA in a manner consistent with a model of regulation by RUST.

Mutants that overexpress *gfp::rsp-6* mRNA affect *rsp-6* expression and splicing in ways that are not entirely consistent with regulation by RUST (Figure 3.3G). The relative expression of *rsp-6(total)* increases 3- to 5-fold in two independent *gfp::rsp-6* overexpression lines ($p<0.05$ for both). However, any changes in the relative expression of *rsp-6(+)* mRNA are not statistically significant, and neither the relative expression nor the proportion of *rsp-6(PTC)* mRNA changes in overexpression strains. As noted above, both *gfp::rsp-6(rEx187)* and *gfp::rsp-6(rEx190)* fail to rescue the sterility of *rsp-6(Δ)* homozygotes and the expression of GFP::RSP-6 fusion proteins appears lower than GFP::RSP-2 or GFP::RSP-4. Thus, I

question whether expressed GFP::RSP-6 fusion proteins are functional or expressed in sufficient quantity to have biological effect.

Rsp-6 mutants do not affect RSP-6 protein abundance

RUST predicts that changes in alternatively spliced mRNAs occur as a homeostatic mechanism to maintain appropriate abundance of a protein. I investigated whether the abundance of RSP-6 changes either in *rsp-6(Δ)/+* deletion heterozygotes or in strains designed to overexpress *gfp::rsp-6*. I measured the relative abundance of RSP-6 with western blots in the same L4-stage strains used for mRNA analysis. Expression of RSP-6 in *rsp-6(Δ)/+* deletion heterozygotes is not significantly different than expression in *rsp-6(+)* (Figure 3.3H). This result confirms the quantification of mRNAs in Figure 3.3F, where the abundance of *rsp-6(+)* mRNA does not change significantly in *rsp-6(Δ)/+* deletion heterozygotes compared to *rsp-6(+)*. Similarly, the abundance of endogenous RSP-6 also does not change in strains designed to overexpress GFP::RSP-6 (Figure 3.3I, lanes 5 and 6 compared to 4). I conclude that endogenous RSP-6 protein levels remain unchanged in *rsp-6(Δ)/+* deletion heterozygotes and in transgenic *gfp::rsp-6* strains.

In summary, patterns of alternative splicing change dramatically in *rsp-6(Δ)/+* deletion heterozygotes, even though the relative expression of RSP-6 protein does not change. I interpret the relative constancy of *rsp-6(+)* mRNA and RSP-6 protein abundance under conditions where the relative expression of *rsp-6(total)* mRNA declines sharply (as in *rsp-6(Δ)/+* deletion heterozygotes) to reflect homeostasis of RSP-6 abundance via RUST.

RSP proteins cross-regulate the splicing patterns of other *rsp* mRNAs

Several lines of evidence from the literature suggest that specific RSP proteins affect splicing patterns of other *rsp* mRNAs. First, RSP orthologues are known regulators of alternative splicing (reviewed in [24]). Second, previous examples of splicing factors regulating the splicing pattern of other splicing factors have been documented, such as the effect of hnRNP L on *hnRNP LL* splicing [162], or of PTB on nPTB splicing patterns [104]. Finally, I demonstrated above that *C. elegans rsp* splicing is altered in *rsp(Δ)/+* deletion heterozygotes. Therefore, I investigated whether individual RSP proteins regulate splicing of *rsp* mRNAs other than their own.

I examined splicing of all six alternatively spliced *rsp* genes in the *rsp* mutants described above. Mutants for these analyses fell into one of two categories: (i) mutants homozygous for one or more *rsp(Δ)* deletions and (ii) strains that overexpress *gfp::rsp-2*, *gfp::rsp-4,* or *gfp::rsp-6* mRNAs. I crossed both categories of mutations into *smg-2(-)* backgrounds. *Rsp-2(Δ)*, *rsp-4(Δ), rsp-5(Δ)* deletion homozygotes are viable and fertile, as are double mutant combinations involving these genes. I grew these strains plus controls in synchronized liquid cultures and harvested L4 stage animals. *Rsp-6(Δ)* deletion homozygotes are sterile. I obtained *rsp-6(Δ)* homozygotes and any double mutant combinations involving *rsp-6(Δ)* by hand picking viable, non-green, L4-stage offspring of *rsp-6(Δ)/nT1[qIs51]* heterozygotes. These samples were compared to handpicked L4-stage *smg-2(-)* animals and handpicked L4-stage *gfp::rsp* transgenic animals in order to minimize any variations in growth conditions.

I quantified *rsp(+)* and *rsp(PTC)* mRNAs of *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7* normalized to *eft-3* in both *rsp(Δ)* and *rsp* overexpression strains. I then calculated the proportions of *rsp(PTC)* mRNAs as a percent of total mRNA for all strains tested and performed statistical tests to determine the significance, if any, of changes in *rsp(PTC)* proportions. I applied a Bonferroni correction to account for the large number of multiple comparisons made in these sets of experiments. The corrected p-values are described in the experiments below. I limit the description to changes in *rsp-5* and *rsp-7* splicing (Figure 3.4) for two reasons. First, changes in *rsp-5* and *rsp-7* splicing patterns are statistically significant after a Bonferroni correction. Second, changes in *rsp-5* and *rsp-7* splicing in *rsp(Δ)* mutants are affected in the opposite manner in *rsp* overexpression mutants. However, all measurements of *rsp* splicing patterns and statistics are shown in Table 3.1 and the original data on which Table 3.1 is based are shown in Table 3.2 (relative expression of *rsp(PTC)* and *rsp(+)* transcripts).

Regulation of *rsp-5* splicing

RSP-2 and RSP-4 affect alternative splicing of *rsp-5* mRNA. *Rsp-2(Δ)* and *rsp-4(Δ)* mutations decrease the proportion of *rsp-5(PTC)* mRNA (from 43% to 15.8% for *rsp-2(Δ)* and from 43% to 28.1% for *rsp-4(Δ)*; p-value 0.0012 and 0.06, respectively; Figure 3.4A). In an *rsp-2(Δ) rsp-4(Δ)* double mutant, the proportion of *rsp-5(PTC)* splicing decreases even further, to 8.4%; p-value 0.0009. The decreased proportion of *rsp-5(PTC)* mRNA in *rsp-2(Δ) rsp-4(Δ)* mutants is statistically significant when compared either to *rsp-2(Δ)* or to *rsp-4(Δ)* single mutants; p-value 0.03 between [*rsp-2(Δ)*] and [*rsp-2(Δ) rsp-4(Δ)*], and p-value 0.006

between [*rsp-4(Δ)*] and [*rsp-2(Δ) rsp-4(Δ)*]. A double *rsp-4(Δ) rsp-6(Δ)* mutant decreases the proportion of *rsp-5(PTC)* from 30.5% to approximately 15% (p-value 0.028).

Overexpression of GFP::RSP-2 or GFP::RSP-4 affects alternative splicing of *rsp-5* in a manner opposite to that of *rsp-2(Δ)* and *rsp-4(Δ)* mutants (Figure 3.4B). Strains harboring *rEx188*, *rEx186*, or *rEx191*, which overexpress either GFP::RSP-2 or GFP::RSP-4 (see Figure 3.2) express an increased proportion of *rsp-5(PTC)* compared to an *rsp(+)* strain ($p<0.05$ for all three comparisons).

The relative expression of *rsp-5(+)* and *rsp-5(PTC)* mRNAs also changes in mutants that delete *rsp-2* or *rsp-4* or overexpress GFP::RSP-2 or GFP::RSP-4 (see Table 3.2). In general, the absolute changes are not statistically significant, but the patterns of changes are consistent, and the proportions of *rsp-5(PTC)* vs. *rsp-5(+)* mRNA are statistically significant. In an *rsp-2(Δ) rsp-4(Δ)* double mutant the relative abundance of *rsp-5(+)* mRNA increases about 3-fold (p-value 0.006), while that of *rsp-5(PTC)* decreases about 2-fold (Table 3.2). Smaller effects are observed in *rsp-2(Δ)* and *rsp-4(Δ)* single mutants. In a strain containing *rEx188*, which overexpresses GFP::RSP-2 (see Table 3.2), expression of *rsp-5(+)* mRNA is possibly decreased, while expression of *rsp-5(PTC)* is apparently unchanged. In strains containing *rEx186* or *rEx191*, which overexpress GFP::RSP-4 (see Table 3.2), expression of *rsp-5(+)* is possibly decreased, while expression of *rsp-5(PTC)* is increased 2- to 2.5-fold.

I conclude that RSP-2 and RSP-4 regulate the alternative splicing of *rsp-5* in a manner that is consistent with cross-regulation by RUST. A surplus of RSP-2 or RSP-4 protein shifts the splicing pattern of *rsp-5* such that an increased proportion of *rsp-5* mRNA is *rsp-5(PTC)* and a decreased proportion is *rsp-5(+)*.

Regulation of *rsp-7* splicing

RSP-4 and RSP-6 affect alternative splicing of *rsp-7* mRNA. *Rsp-4(Δ)* and *rsp-6(Δ)* mutations decrease the proportion of *rsp-7(PTC)* mRNA (from 54% to ~33% for an *rsp-4(Δ)* single and an *rsp-4(Δ) rsp-2(Δ)* double mutant, and from 46.1% to 34.8% for an *rsp-6(Δ)* single mutant; $p<0.01$ for *rsp-6(Δ)*; Figure 3.4C). The changes in *rsp-7(PTC)* proportion are not significant in an *rsp-4(Δ)* single mutant after a Bonferroni correction. However, in an *rsp-4(Δ);rsp-6(Δ)* double mutant, the proportion of *rsp-7(PTC)* mRNA decreases even further (to 21.8%, $p<0.05$ compared to an *rsp-6(Δ)* single deletion). Additionally, overexpression of GFP::RSP-4 affects alternative splicing of *rsp-7* in a manner that is opposite to that of an *rsp-4(Δ)* deletion (Figure 3.4D). In two independent GFP::RSP-4 overexpression strains, the proportion of *rsp-7(PTC)* mRNA increases significantly ($p<0.05$ for *rEx186*; $p=0.3$ for *rEx191*).

The relative expression of *rsp-7(+)* and *rsp-7(PTC)* mRNAs change in several *rsp* mutants (Table 3.2). In general, these changes are not statistically significant, though the patterns of changes are consistent. Expression of *rsp-7(+)* increases in the *rsp-4(Δ)* single mutant and in the *rsp-4(Δ) rsp-2(Δ)* double mutant, while expression of *rsp-7(PTC)* may decrease slightly. Expression of *rsp-7(PTC)* decreases in *rsp-6(Δ)* mutants, while expression of *rsp-7(+)* is unchanged. In an *rsp-4(Δ);rsp-6(Δ)* double mutant, expression of *rsp-7(PTC)* decreases approximately 3-fold (p-value 0.002). In strains containing *rEx186* and *rEx191*, which overexpress GFP::RSP-4, expression of *rsp-7(+)* decreases about 2-fold, while expression of *rsp-7(PTC)* mRNA changes little.

I conclude that RSP-4 and RSP-6 regulate alternative splicing of *rsp-7* in a manner that is consistent with cross-regulation by RUST. A surplus of RSP-4 protein shifts the splicing pattern of *rsp-7* such that an increased proportion of *rsp-7* mRNA is *rsp-7(PTC)* and a decreased proportion is *rsp-7(+)*. Conversely, a deficit of RSP-4 or RSP-6 has the opposite effect.

Discussion

I built groups of mutants that fail to express, reduce, or overexpress RSP proteins. Using these mutants, I tested whether *rsp* splicing changes in patterns predicted by the RUST model. RUST predicts splicing patterns of mRNAs change to maintain protein homeostasis. Changes in splicing patterns increase or decrease the proportion of mRNA translated into functional protein.

I tested RUST in two ways. First, I tested if reducing or overexpressing a single RSP protein affects the splicing pattern of its own mRNA. Second, I tested if one or more RSP protein affects the splicing of other *rsp* mRNAs. My experiments uncovered regulation of *rsp* mRNA splicing consistent with a model of RUST. However, the mechanism by which RSP proteins regulate their own and each other's splicing is unclear.

Feedback regulation of *rsp* splicing

I observed changes in *rsp-2, rsp-4, rsp-5,* and *rsp-6* mRNA expression in heterozygous *rsp-2(Δ)*, *rsp-4(Δ)*, *rsp-5(Δ)*, and *rsp-6(Δ)* strains, respectively (Figure 3.3A, C, D, F). The proportion of *rsp(PTC)* mRNA decreases in *rsp(Δ)* heterozygotes. *Rsp(Δ)* heterozygotes are designed to reduce expression of RSP proteins. RUST predicts a decrease

in RSP protein should decrease the proportion of *rsp(PTC)* mRNA so the proportion of *rsp(+)* mRNA increases and is translated into functional protein to achieve protein homeostasis. Therefore, these data support a model of RUST.

Changes in *rsp-4* splicing also occur in two lines that overexpress RSP-4 protein (Figure 3.3D). The relative expression of *rsp-4(PTC)* increases in strains where RSP-4 is overexpressed. These changes are consistent with a model of RUST wherein, the presence of excess RSP-4 protein increases the relative expression of *rsp-4(PTC)* mRNA. Lines designed to overexpress RSP-2 and RSP-6 proteins show no changes in the proportion of *rsp-2(PTC)* or *rsp-6(PTC)*, respectively (Figure 3.3B and G). These data suggest *rsp-2* and *rsp-6* splicing patterns are unaffected by the overexpression of RSP-2 and RSP-6 proteins, respectively, and do not support a model of RUST. However, it is unclear why transgenic strains fail to affect *rsp-2* or *rsp-6* splicing patterns. The simplest explanation is neither RSP-2 nor RSP-6 transgenes function like endogenous RSP proteins.

RUST predicts changes in splicing maintain *rsp(+)* mRNA levels so a consistent level of RSP protein is translated. Using an RSP-6 antibody, I observed no significant changes in RSP-6 protein levels in strains heterozygous for *rsp-6(Δ)* (Figure 3.3H). These data support a model where changes in the splicing of *rsp-6* mRNA occur so *rsp-6(+)* mRNA is constant and maintains protein homeostasis.

How might RSP proteins regulate the splicing of their own mRNAs? Perhaps the simplest explanation is a direct interaction between RSP proteins and their own pre-mRNAs. Mammalian SR and hnRNP proteins are known to associate with their own pre-mRNAs to regulate the splicing of their own mRNAs through RUST. For example, the mammalian

orthologue of RSP-6, SRSF3, was recently shown to bind its own pre-mRNA [50]. hnRNP L is also known to associate with its own mRNA to direct the splicing of a PTC-containing exon [162]. SR proteins are known regulators of alternative splicing, therefore RSP proteins might associate with their own pre-mRNAs to mark alternative exons to be included in mature mRNAs. Though such an explanation seems likely to explain *rsp* splicing patterns, other explanations could also explain changes in *rsp* splicing.

SRSF3 (orthologue of RSP-6) and other SR proteins are known to interact with RNA polymerase II [73]. *In vitro*, interactions between SR proteins and Pol II are required to efficiently transcribe and splice reporter mRNAs. Perhaps RSP proteins interact with their own pre-mRNAs as they are being transcribed. RSP proteins could therefore partly regulate the transcription and the splicing of their own pre-mRNAs.

Other unknown factors might regulate the splicing of *rsp* mRNAs in response to overexpression or reduced expression of RSP proteins. RSP proteins contain RS domains that are important for protein-protein interactions with components of the splicing machinery [24]. RS domains are also present in a large number of SR-like proteins that have functions in splicing [14]. Perhaps other splicing factors interact with RSP proteins through their RS domains and regulate *rsp* splicing patterns in response to a reduction or overexpression of RSP proteins.

Cross-regulation of *rsp* splicing patterns

I observed the splicing patterns of *rsp-5* and *rsp-7* are affected in mutants that fail to express RSP proteins or overexpress RSP proteins (Figure 3.4). Specifically, RSP-2 and

RSP-4 both affect the splicing pattern of *rsp-5*. Single *rsp-2(Δ)* and *rsp-4(Δ)* deletions each reduce the proportion of *rsp-5(PTC)* mRNA. A double *rsp-2(Δ) rsp-4(Δ)* deletion mutant further decreases the proportion of *rsp-5(PTC)* mRNA. Strains that overexpress RSP-2 or RSP-4 affect *rsp-5* splicing in the opposite pattern by increasing the proportion of *rsp-5(PTC)* mRNA. RSP-4 affects the splicing pattern of *rsp-7*. Single *rsp-4(Δ)* mutants, and double [*rsp-2(Δ) rsp-4(Δ)*] or [*rsp-4(Δ);rsp-6(Δ)*] all decrease the proportion of *rsp-7(PTC)* mRNA. Overexpression of RSP-4 affects *rsp-7* in the opposite manner by increasing the proportion of *rsp-7(PTC)* mRNA. These data show RSP-2 and RSP-4 cross-regulate *rsp-5* splicing and RSP-4 also cross-regulates *rsp-7* splicing.

How RSP proteins cross-regulate other *rsp* mRNAs, and if such regulation is direct or indirect is unknown. The possible mechanisms of RSP cross-regulation could be similar or the same as RSP autoregulation. Direct interaction between RSP proteins and other *rsp* pre-mRNAs seems the most likely explanation. As previously described, SRSF3 (mammalian RSP-6) binds its own pre-mRNA. SRSF3 also binds *SRSF2* (orthologous to *rsp-4* and *rsp-5*), *SRSF5* (orthologous to *rsp-2*), and *SRSF7* (similar to *rsp-6*) pre-mRNAs [50]. hnRNP L is known to cross-regulate splicing of *hnRNP LL* mRNA [162]. Though the mechanism by which hnRNP L regulates *hnRNP LL* splicing is not known, a direct interaction seems likely. hnRNP L autoregulates its own splicing by binding to a set of CA clusters in the intron adjacent to exon 6A of *hnRNP L* mRNA. Those CA clusters are also present adjacent to a similar exon 6A in *hnRNP LL* mRNA, suggesting hnRNP L directly associates with *hnRNP LL* mRNA at similar nucleotides, and directs *hnRNP LL* splicing. Other explanations described as mechanisms for RSP autoregulation could also explain cross-regulation. These

include RSP proteins affecting the transcription of different *rsp* pre-mRNAs or other unknown factors interacting with *rsp* pre-mRNAs to direct *rsp* splicing patterns.

Why might RSP proteins regulate the splicing patterns of other *rsps*? One explanation is that RSP proteins might have redundant roles in *C. elegans* and the changes in splicing I observed are to compensate for the loss or overexpression of one or more other RSP proteins. Too much of RSP-2 protein might be compensated for by shifting the splicing pattern to increase the proportion of *rsp-5(PTC)* mRNA, turning down expression of RSP-5 protein. Conversely, too little of RSP-2 protein and *rsp-5* splicing can be changed to increase *rsp-5(+)* proportion and RSP-5 protein expression. Several lines of evidence suggest RSP/SR proteins may be redundant. First, in mammalian cell culture, each SR protein can restore splicing activity in cytoplasmic S100 fractions [19]. Second, with the exception of the *rsp-6(Δ)* homozygotes, the single *rsp(Δ)* deletion mutants described in this chapter have no visible phenotypes. Double *rsp-2(Δ);rsp-6(Δ)* mutants appear synthetic lethal, but other double *rsp(Δ)* deletion homozygotes have no visible phenotypes. Finally, published experiments show RNAi against any single *rsp* gene with the exception of *rsp-3* show no visible phenotype [41,42].

Another explanation might be specific RSP proteins are needed in specific tissues or at specific developmental time points. By changing the splicing patterns of *rsp* mRNAs, the expression of RSP proteins could be affected in tissue-specific or developmentally regulated patterns. For example, PTB and nPTB are hnRNP proteins whose pre-mRNAs are both alternatively spliced to generate PTC and + isoforms [104]. PTB cross-regulates *nPTB* splicing to generate *nPTB(PTC)* isoforms in all but neuronal tissues. Similar cross-regulation

by RSP proteins could affect the splicing of other *rsp* mRNAs. It is also possible RSP proteins could cross-regulate the splicing of other *rsp* mRNAs at specific developmental time points. In chapter 2, I demonstrated the splicing patterns of four *rsp* mRNAs are under tight developmental control. Perhaps cross-regulation by RSP proteins plays roles in regulating developmental patterns of *rsp* alternative splicing. It is also possible the regulation of *rsp-5* and *rsp-7* splicing I observed in L4-stage animals is specific to this developmental time point and other RSP proteins might regulate other *rsp* mRNAs at other developmental time points. Measuring *rsp* splicing patterns in *rsp* mutants at other developmental time points might uncover additional examples of RSP cross-regulation.

Figure 3.1. Maps of *rsp* deletion alleles.

I obtained four *rsp* deletion alleles that delete part or most of an *rsp*. Deletion alleles of *rsp-2* and *rsp-4* delete regions including the AUG translation initiation codon, and deletion alleles of *rsp-5* and *rsp-6* delete nearly the entirety of each respective gene. A line underneath each gene represents the deleted region and is indicated by the allele number for the *rsp* deletion. The splicing patterns for each *rsp* are also represented. Red exons represent the alternatively spliced exons that introduce PTCs. Primers were designed to allow multiplexing PCR to detect homozygous wild-type, homozygous deletion, or heterozygous alleles in the same PCR reaction and are indicated as “for,” “rev wt,” and “rev mut.”

Figure 3.1

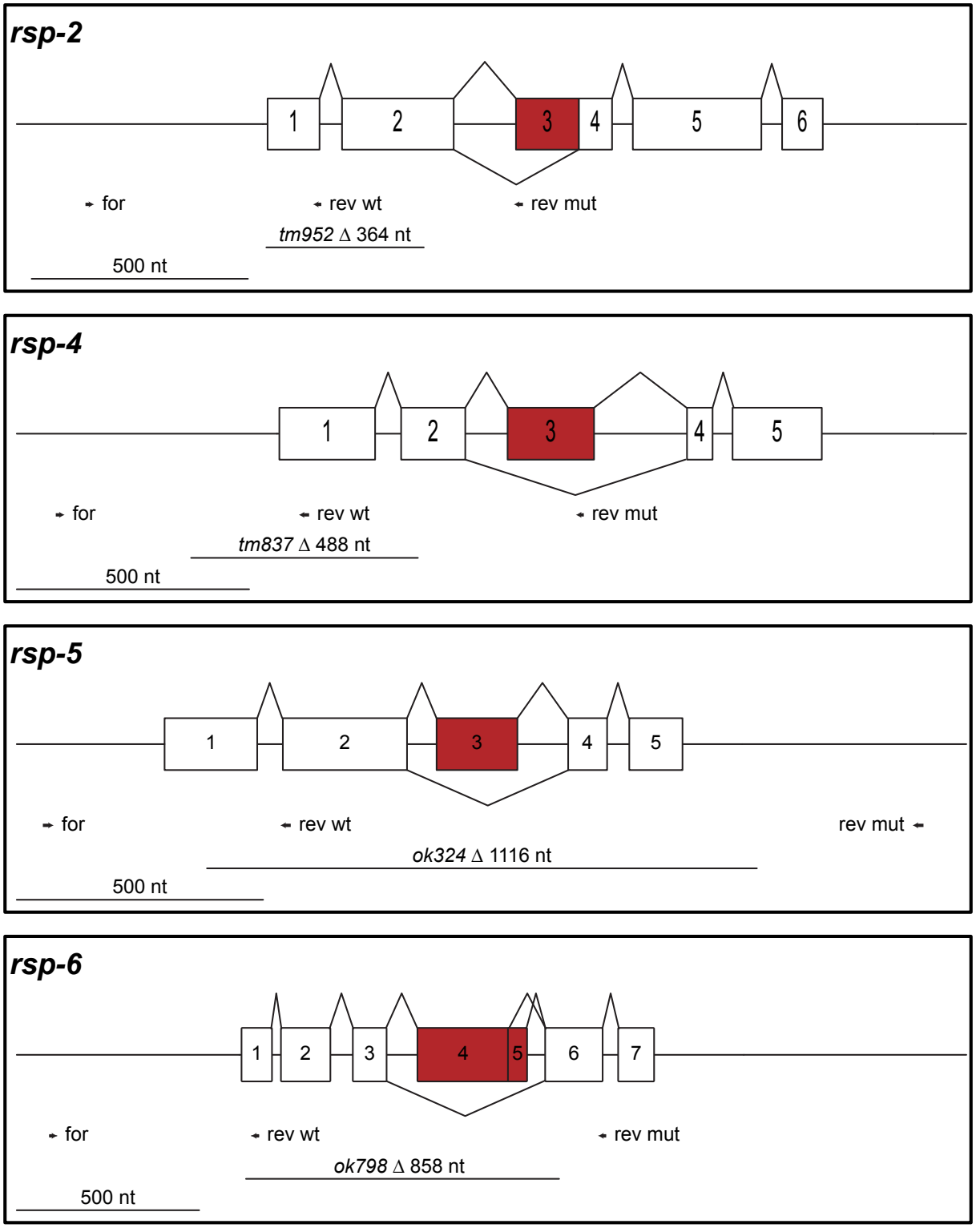

rsp-2
1
2
3
4
5
6
for
rev wt
rev mut
tm952 Δ 364 nt
500 nt
rsp-4
1
2
3
4
5
for
rev wt
rev mut
tm837 Δ 488 nt
500 nt
rsp-5
1
2
3
4
5
for
rev wt
rev mut
ok324 Δ 1116 nt
500 nt
rsp-6
1
2
3
4
5
6
7
for
rev wt
rev mut
ok798 Δ 858 nt
500 nt

Figure 3.2. Transgenic *gfp::rsp* alleles express GFP::RSP fusion proteins.

A) Map of *gfp::rsp-2, gfp::rsp-4,* and *gfp::rsp-6* extrachromosomal arrays. *Rsp* genes were cloned from genomic DNA (top of each *rsp* map with the splicing patterns for each *rsp* shown). I then removed the alternatively spliced exons and fused the flanking exons together. Finally, I cloned *rsp(+)* forms into a vector that fused *gfp* and *rsp(+)* isoforms (bottom of each *rsp* map) and drives expression with the *eft-3* promoter. *Unc-54* 3' UTR distinguishes exogenous *rsp(+)* mRNAs from endogenous. B) An anti-GFP western blot detects a strain expressing free GFP as a control (indicated GFP positive control) and GFP::RSP fusion proteins in *gfp::rsp-2, gfp::rsp-4,* and *gfp::rsp-6* transgenic strains. Green arrows represent GFP and GFP::RSP fusion proteins.

Figure 3.2

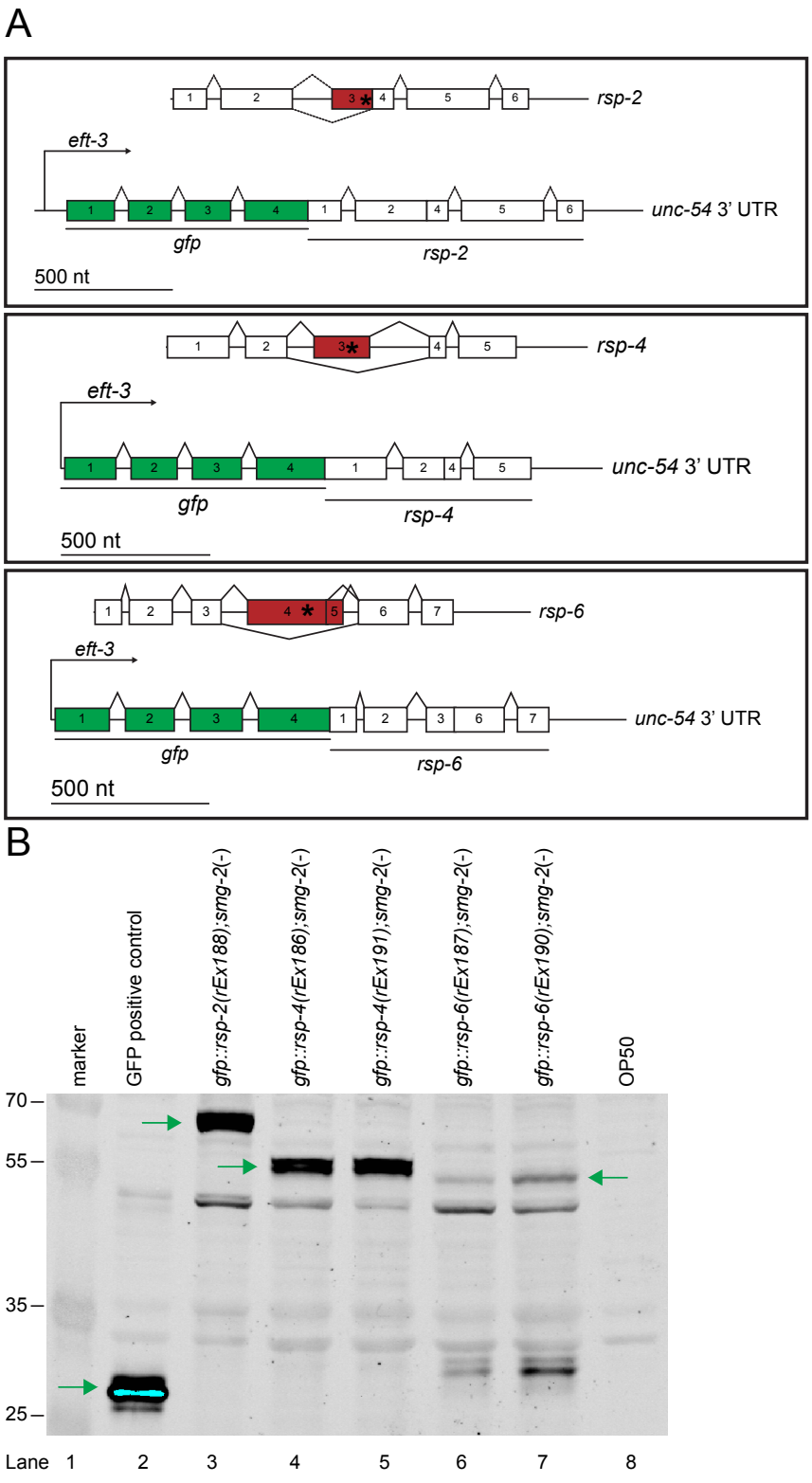

A
rsp-2
eft-3
unc-54 3' UTR
gfp
rsp-2
500 nt
rsp-4
eft-3
unc-54 3' UTR
gfp
rsp-4
500 nt
rsp-6
eft-3
unc-54 3' UTR
gfp
rsp-6
500 nt
B
marker
GFP positive control
gfp::rsp-2(rEx188);smg-2(-)
gfp::rsp-4(rEx186);smg-2(-)
gfp::rsp-4(rEx191);smg-2(-)
gfp::rsp-6(rEx187);smg-2(-)
gfp::rsp-6(rEx190);smg-2(-)
OP50
70
55
35
25
Lane 1 2 3 4 5 6 7 8

Figure 3.3. The pre-mRNA splicing patterns of *rsp-2, rsp-4, rsp-5,* and *rsp-6* are autoregulated.

In L4-stage samples, I measured *rsp-2* (A), *rsp-4* (C), *rsp-5* (E), and *rsp-6* (F) mRNAs in *rsp-2(Δ), rsp-4(Δ), rsp-5(Δ),* and *rsp-6(Δ)* deletion heterozygotes (white bars), respectively compared to controls (black bars). **All samples are also *smg-2(-).*** In general, the relative expression of *rsp(total)* and *rsp(PTC)* mRNA levels decrease while *rsp(+)* mRNAs are unaffected (left charts). The relative expression of *rsp* mRNAs is normalized to *eft-3*. The proportion of *rsp(PTC)* [*rsp(PTC)*/sum of *rsp(PTC)* and *rsp(+)*] decreases in mutants heterozygous for an *rsp* compared to a control (right charts).

In L4-stage samples, I measured *rsp-2* (B), *rsp-4* (D), and *rsp-6* (G) mRNAs in *gfp::rsp-2, gfp::rsp-4,* and *gfp::rsp-6* strains (blue and green bars), respectively compared to controls (black bars). **All samples are also *smg-2(-).*** All measurements are of endogenous *rsp* mRNAs except for *rsp(total)* in transgenic strains that includes both endogenous and exogenous *rsp* mRNA. The left charts plot the relative expression of *rsp(PTC), rsp(+),* and *rsp(total)* mRNAs normalized to *eft-3*. The far right charts plot the proportion of *rsp(PTC)* [*rsp(PTC)*/sum of *rsp(PTC)* and *rsp(+)*]. Compared to controls, *rsp-2* and *rsp-6* mRNAs show no statistical changes in the relative expression of PTC or + mRNAs, or in the proportion of PTC mRNA. In *gfp::rsp-4* strains, I observe an increase in the relative expression of *rsp-4(total)* mRNA and *rsp-4(PTC)* mRNA while endogenous *rsp-4(+)* mRNA levels are unaffected.

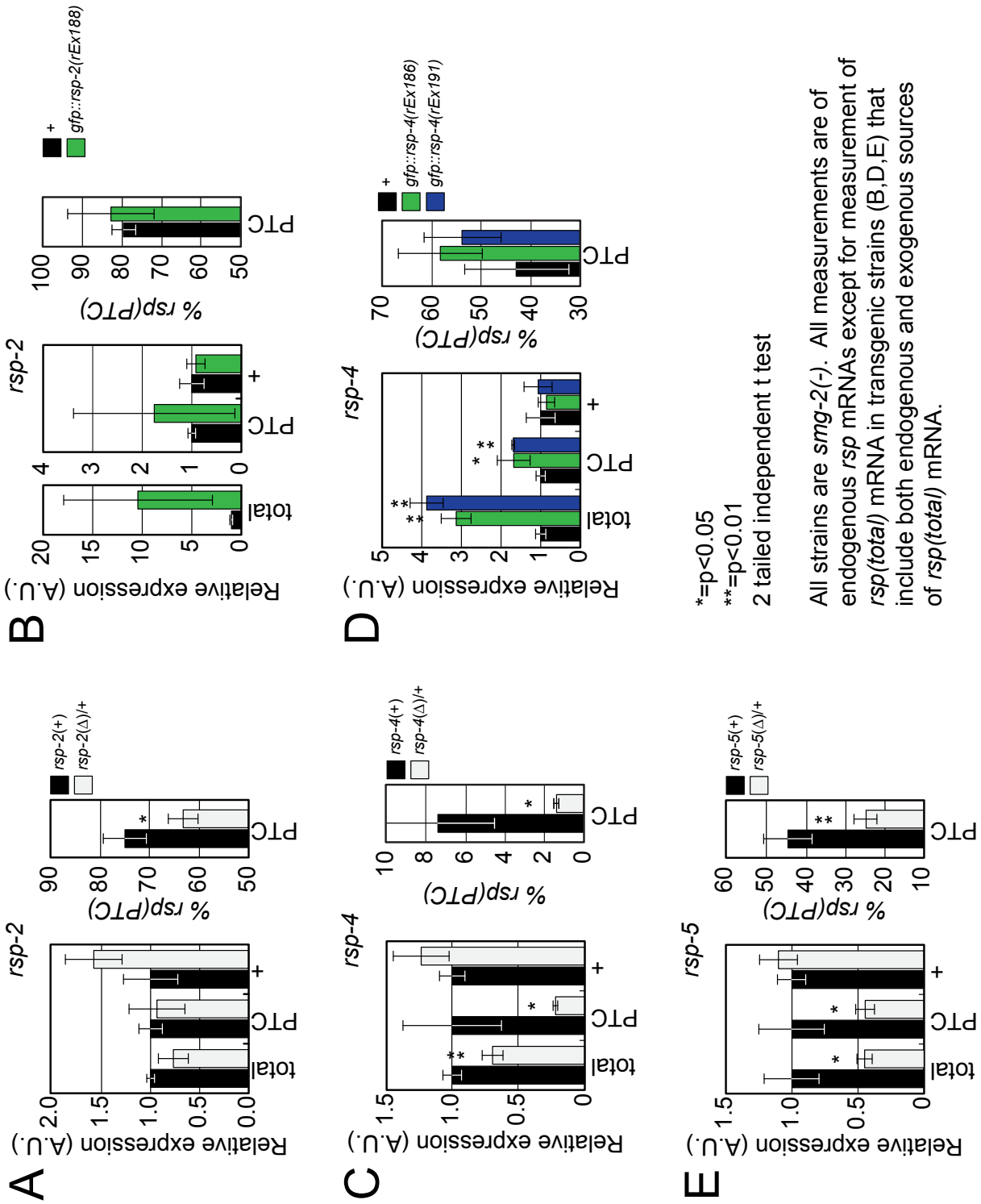
Figure 3.3. The pre-mRNA splicing patterns of *rsp-2, rsp-4, rsp-5,* and *rsp-6* are autoregulated.

I quantified endogenous RSP-6 protein levels normalized to actin in L4-stage heterozygous *rsp-6(Δ)* mutants (H) and L4-stage strains overexpressing *gfp::rsp-6* (I). **All samples are also *smg-2(-)*.** Neither *rsp-6(Δ)* heterozygotes, nor *gfp::rsp-6* strains show differences in the abundance of endogenous RSP-6 (Lanes 1 and 2 in 3.3H and lanes 4-6 in 3.3I). *Gfp::rsp-6* transgenic strains were also probed with anti-GFP and show expression of the fusion GFP:RSP-6 protein.

Adult samples in western blots are controls to show the specificity of the RSP-6 antibody.

For A-I, n=3 and error bars in A-G indicate +/- SD.

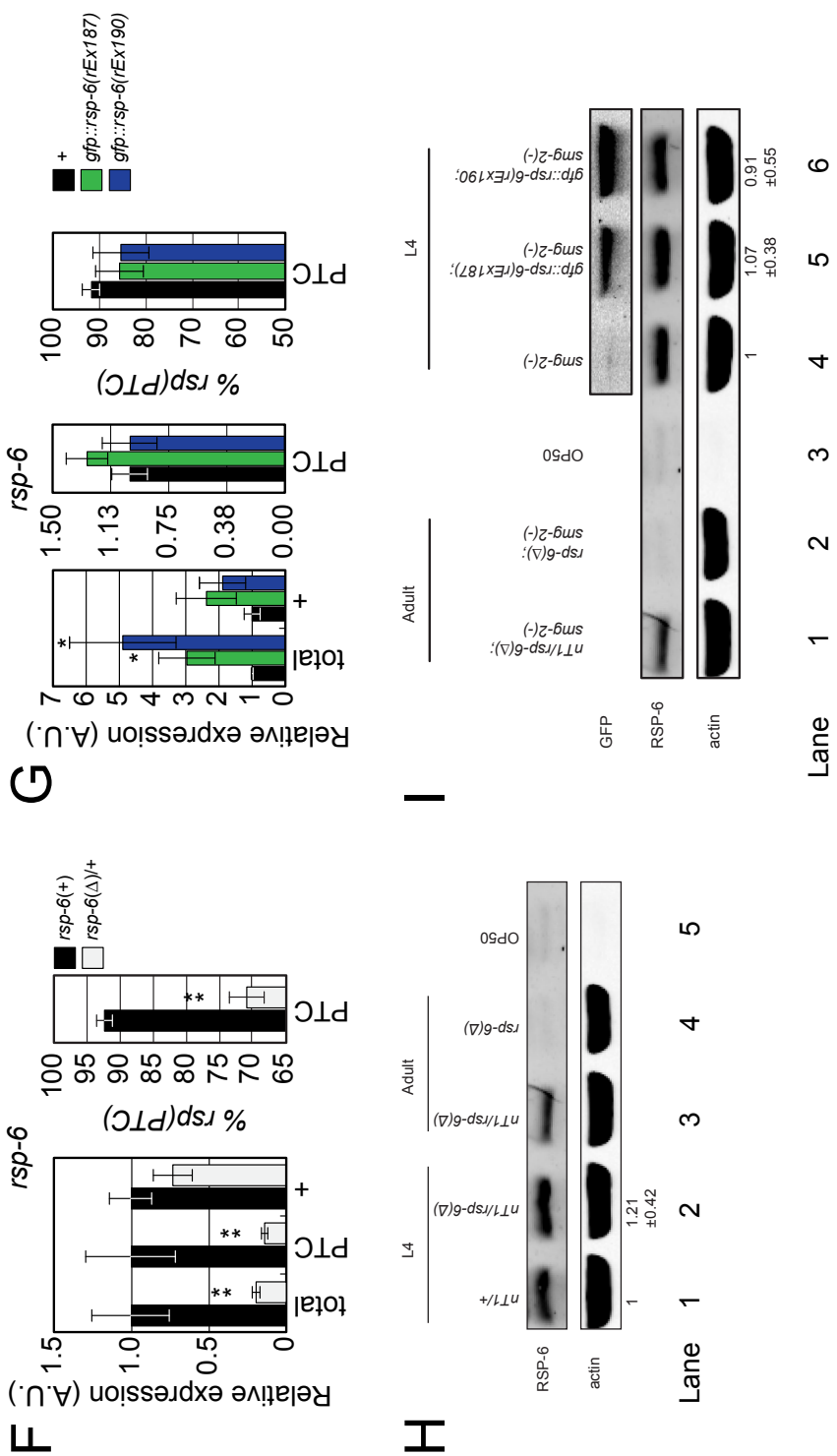
Figure 3.3



*=p<0.05
**=p<0.01
2 tailed independent t test

All strains are *smg-2(-)*. All measurements are of endogenous *rsp* mRNAs except for measurement of *rsp(total)* mRNA in transgenic strains (B,D,E) that include both endogenous and exogenous sources of *rsp(total)* mRNA.

Figure 3.3 Continued



*=p<0.05
**=p<0.01
2 tailed independent t test

All strains are *smg-2(-)*. All measurements are of endogenous *rsp* mRNAs except for measurement of *rsp(total)* mRNA in transgenic strains (B,D,E) that include both endogenous and exogenous sources of *rsp(total)* mRNA.

Figure 3.4. RSP proteins cross-regulate *rsp-5* and *rsp-7* splicing patterns.

A) Removal of RSP-2 or RSP-4 decreases the proportion of *rsp-5(PTC)* mRNA compared to a control strain (*rsp* +). This effect is additive as a double *rsp-2(Δ) rsp-4(Δ)* mutant further decreases the proportion of *rsp-5(PTC)*.

B) Overexpression of RSP-2 or RSP-4 increases the proportion of *rsp-5(PTC)* compared to a control strain (*rsp* +).

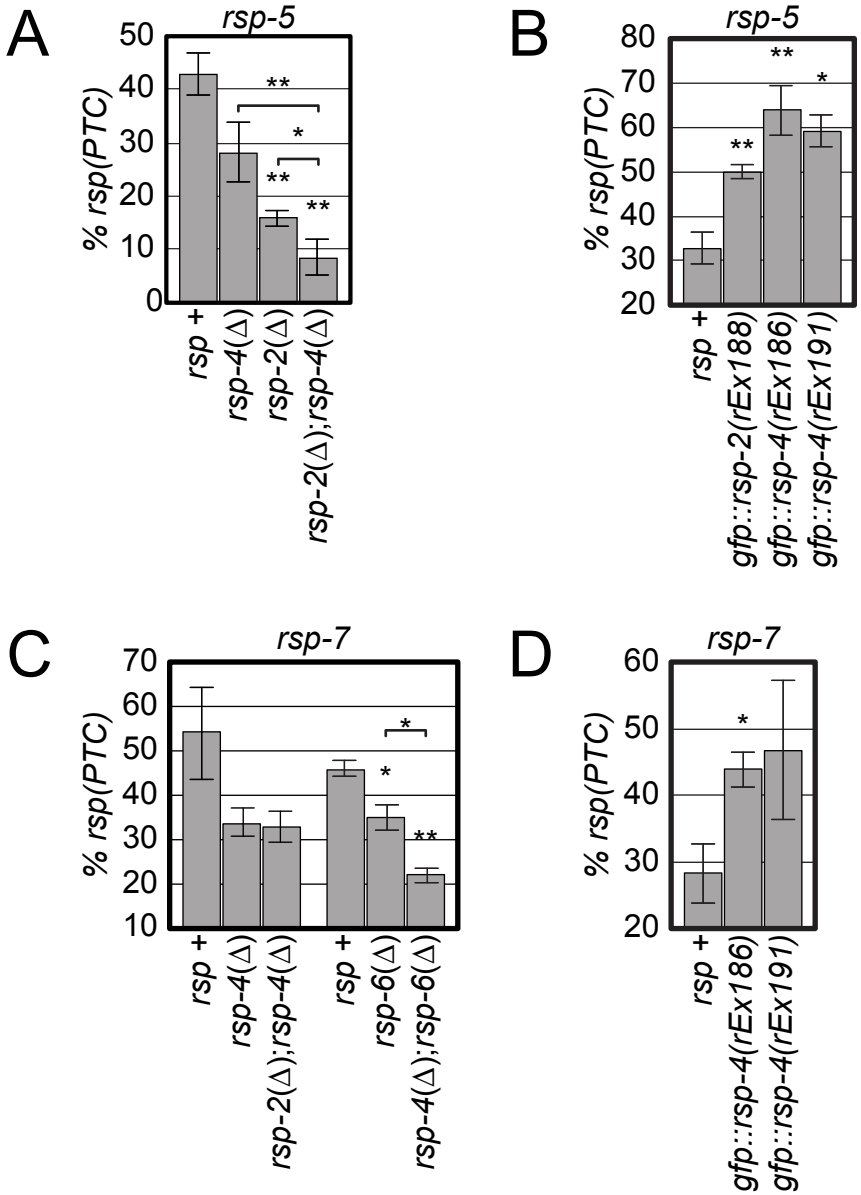
C) Removal of RSP-4 or RSP-6 decreases the proportion of *rsp-7(PTC)* compared to a control strain (*rsp* +).

D) Overexpression of RSP-4 increases the proportion of *rsp-7(PTC)* compared to a control strain (*rsp* +). Overexpression of RSP-6 does not affect *rsp-7(PTC)* proportion.

All samples are *smg-2(-)*.

The proportion of *rsp(PTC)* [*rsp(PTC)*/sum of *rsp(PTC)* and *rsp(+)*] is plotted in each chart. I performed a t-test for comparisons between a control (*rsp* +) and *rsp(Δ)* deletion homozygotes or strains overexpressing *gfp::rsp* mRNAs. I corrected for multiple tests with a Bonferroni correction. Corrected p-values are indicated as < 0.05 or < 0.01. Error bars indicate +/- SD, n = 3.

Figure 3.4



All samples are *smg-2(-)*

P-values after Bonferroni correction
*=p<0.05
**=p<0.01
2 tailed independent t test

Table 3.1. Measurements of *rsp* splicing patterns in all *rsp* mutants used in cross-regulation analysis.

I measured the proportion of *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7* (PTC) [*rsp(PTC)/*sum of *rsp(PTC)* and *rsp(+)*] in nearly all *rsp(Δ)* single, *rsp (Δ) rsp (Δ)* double, and *rsp* overexpression mutants described in this chapter. **All samples are also *smg-2(-).*** As described in Figure 3.5, I performed a t-test to compare control strains *(rsp +)* and *rsp* mutant strains (either *rsp(*Δ*)* or overexpressing *gfp::rsp* mRNA) and corrected for multiple tests with a Bonferroni correction. Both an uncorrected and corrected p-value are shown. All strains in Table 3.1 are *smg-2(-)*. IC = incomplete measurement. For each measurement n = 2 or 3 samples. These tables do not show the measurements of *rsp(PTC)* on the same transgenic strain. For example, *rsp-4(PTC)* proportion is not plotted in *gfp::rsp-4* overexpression strains. Those data are shown in Figure 3.4 and address autoregulation.

Table 3.1

rsp-1

Strain	% PTC ± SD	P-value	Corrected P-value
rsp(+)	67.4 ± 1.7	-	-
rsp-2(Δ)	80.0 ± 8.1	0.06	0.36
rsp-4(Δ)	65.2 ± 14.7	0.81	1
rsp-5(Δ)	78.7 ± 5.8	0.03	0.18
rsp-2(Δ) rsp-4(Δ)	59.9 ± 2.0	0.0078	0.0468
rsp-2(Δ) rsp-5(Δ)	74.9 ± 7.7	0.18	1
rsp-4(Δ) rsp-5(Δ)	69.4 ± 8.4	0.72	1
rsp(+)	67.0 ± 1.3	-	-
rsp-6(Δ)	76.7 ± 3.3	0.009	0.027
rsp-4(Δ);rsp-6(Δ)	72.9 ± 0.9	0.012	0.036
rsp-5(Δ);rsp-6(Δ)	89.9 ± 3.2	0.001	0.003
rsp(+)	74.5 ± 3.8	-	-
rsp-2(rEx188)	83.9 ± 3.6	0.04	0.12
rsp-4(rEx186)	IC	-	-
rsp-4(rEx191)	IC	-	-
rsp-6(rEx187)	71.8 ± 8.3	0.63	1
rsp-6(rEx190)	68.5 ± 11.2	0.43	1

rsp-2

Strain	% PTC ± SD	P-value	Corrected P-value
rsp(+)	71.2 ± 2.6	-	-
rsp-2(Δ)	-	-	-
rsp-4(Δ)	77.3 ± 1.2	0.02	0.06
rsp-5(Δ)	74.0 ± 4.6	0.41	1
rsp-2(Δ) rsp-4(Δ)	-	-	-
rsp-2(Δ) rsp-5(Δ)	-	-	-
rsp-4(Δ) rsp-5(Δ)	77.4 ± 3.8	0.07	0.21
rsp(+)	61.6 ± 11.7	-	-
rsp-6(Δ)	63.9 ± 2.1	0.75	1
rsp-4(Δ);rsp-6(Δ)	67.2 ± 4.2	0.85	1
rsp-5(Δ);rsp-6(Δ)	63.6 ± 7.7	0.58	1
rsp(+)	63.7 ± 4.9	-	-
rsp-2(rEx188)	-	-	-
rsp-4(rEx186)	77.0 ± 1.3	0.01	0.04
rsp-4(rEx191)	71.4 ± 0.9	0.13	0.52
rsp-6(rEx187)	58.9 ± 7.3	0.4	1
rsp-6(rEx190)	59.8 ± 13.5	0.66	1

rsp-4

Strain	% PTC ± SD	P-value	Corrected P-value
rsp(+)	15.1 ± 1.2	-	-
rsp-2(Δ)	19.7 ± 7.5	0.35	1
rsp-4(Δ)	-	-	-
rsp-5(Δ)	14.5 ± 6.0	0.88	1
rsp-2(Δ) rsp-4(Δ)	-	-	-
rsp-2(Δ) rsp-5(Δ)	14.2 ± 5.3	0.8	1
rsp-4(Δ) rsp-5(Δ)	-	-	-
rsp(+)	12.4 ± 1.4	-	-
rsp-6(Δ)	14.6 ± 1.6	0.16	0.32
rsp-4(Δ);rsp-6(Δ)	-	-	-
rsp-5(Δ);rsp-6(Δ)	21.1 ± 3.5	0.03	0.06
rsp(+)	6.1 ± 1.0	-	-
rsp-2(rEx188)	8.1 ± 0.33	0.03	0.09
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	5.1 ± 2.0	0.48	1
rsp-6(rEx190)	5.3 ± 3.3	0.7	1

rsp-5

Strain	% PTC ± SD	P-value	Corrected P-value
rsp(+)	43.0 ± 4.0	-	-
rsp-2(Δ)	15.8 ± 1.5	0.0004	0.0012
rsp-4(Δ)	28.1 ± 5.6	0.02	0.06
rsp-5(Δ)	-	-	-
rsp-2(Δ) rsp-4(Δ)	8.4 ± 3.4	0.0003	0.0009
rsp-2(Δ) rsp-5(Δ)	-	-	-
rsp-4(Δ) rsp-5(Δ)	-	-	-
rsp(+)	30.5 ± 3.9	-	-
rsp-6(Δ)	24.4 ± 2.2	0.0778	0.1556
rsp-4(Δ);rsp-6(Δ)	15.4 ± 0.4	0.0142	0.0284
rsp-5(Δ);rsp-6(Δ)	-	-	-
rsp(+)	32.8 ± 3.5	-	-
rsp-2(rEx188)	50.0 ± 1.6	0.002	0.01
rsp-4(rEx186)	63.9 ± 5.6	0.001	0.005
rsp-4(rEx191)	59.2 ± 3.7	0.004	0.02
rsp-6(rEx187)	31.9 ± 2.2	0.77	1
rsp-6(rEx190)	27.0 ± 15	0.55	1

Table 3.1 Continued.

rsp-6

Strain	% PTC ± SD	P-value	Corrected P-value
rsp(+)	92.7 ± 0.3	-	-
rsp-2(Δ)	96.0 ± 2.6	0.02	0.12
rsp-4(Δ)	91.1 ± 0.7	0.09	0.54
rsp-5(Δ)	92.3 ± 1.5	0.66	1
rsp-2(Δ) rsp-4(Δ)	92.3 ± 1.0	0.19	1
rsp-2(Δ) rsp-5(Δ)	94.0 ± 1.4	0.52	1
rsp-4(Δ) rsp-5(Δ)	89.5 ± 4.2	0.27	1
rsp(+)	94.4 ± 0.6	-	-
rsp-2(rEx188)	92.7 ± 0.3	0.01	0.01
rsp-4(rEx186)	IC	-	-
rsp-4(rEx191)	IC	-	-
rsp-6(rEx187)	IC	-	-
rsp-6(rEx190)	IC	-	-

rsp-7

Strain	% PTC ± SD	P-value	Corrected P-value
rsp(+)	54.1 ± 10.6	-	-
rsp-2(Δ)	66.8 ± 10.0	0.21	1
rsp-4(Δ)	33.7 ± 3.0	0.03	0.18
rsp-5(Δ)	57.2 ± 0.9	0.65	1
rsp-2(Δ) rsp-4(Δ)	33.0 ± 3.7	0.03	0.18
rsp-2(Δ) rsp-5(Δ)	65.1 ± 7.1	0.21	1
rsp-4(Δ) rsp-5(Δ)	41.7 ± 4.1	0.13	1
rsp(+)	46.1 ± 1.9	-	-
rsp-6(Δ)	34.8 ± 3.1	0.006	0.018
rsp-4(Δ);rsp-6(Δ)	21.8 ± 1.7	0.0007	0.0021
rsp-5(Δ);rsp-6(Δ)	50.7 ± 9.1	0.43	1
rsp(+)	28.3 ± 4.4	-	-
rsp-2(rEx188)	30.5 ± 2.1	0.48	1
rsp-4(rEx186)	43.8 ± 2.6	0.006	0.03
rsp-4(rEx191)	46.8 ± 10.4	0.06	0.3
rsp-6(rEx187)	23.1 ± 3.5	0.27	1
rsp-6(rEx190)	21. 3 ± 7.9	0.26	1

Table 3.2. Measurements of *rsp* mRNA expression in all *rsp* mutants used in cross-regulation analysis.

I measured the relative expression of *rsp-1, rsp-2, rsp-4, rsp-5, rsp-6,* and *rsp-7* PTC (left tables) and + (right tables) mRNAs normalized to *eft-3* in nearly all *rsp(Δ)* single, *rsp (Δ) rsp (Δ)* double, and *rsp* overexpression mutants described in this chapter. **All samples are also *smg-2(-).*** As described in Figure 3.5, I performed a t-test to compare control strains *(rsp +)* and *rsp* mutant strains and corrected for multiple tests with a Bonferroni correction. Both an uncorrected and corrected p-value are shown. All strains in Table 3.2 are *smg-2(-)*. IC = incomplete measurement. For each measurement n = 2 or 3 samples. These tables do not show the measurements of *rsp(PTC)* on the same transgenic strain. For example, *rsp-4(PTC)* proportion is not plotted in *gfp::rsp-4* overexpression strains. Those data are shown in Figure 3.4 and address autoregulation.

Table 3.2

rsp(PTC)

rsp-1

Strain	Rel. Exp. (PTC) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.59	-	-
rsp-2(Δ)	1.32 ± 0.53	0.43	1
rsp-4(Δ)	1.07 ± 0.61	0.87	1
rsp-5(Δ)	1.62 ± 0.26	0.15	0.93
rsp-2(Δ) rsp-4(Δ)	1.05 ± 0.33	0.89	1
rsp-2(Δ) rsp-5(Δ)	1.64 ± 0.11	0.14	0.83
rsp-4(Δ) rsp-5(Δ)	1.48 ± 0.23	0.3	1
rsp(+)	1 ± 0.24	-	-
rsp-6(Δ)	1.04 ± 0.23	0.39	1
rsp-4(Δ);rsp-6(Δ)	0.82 ± 0.22	0.13	0.38
rsp-5(Δ);rsp-6(Δ)	0.98 ± 0.34	0.76	1
rsp(+)	1 ± 0.19	-	-
rsp-2(rEx188)	1.05 ± 0.15	0.76	1
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	1.19 ± 0.16	0.26	0.79
rsp-6(rEx190)	0.99 ± 0.13	0.95	1

rsp-2

Strain	Rel. Exp. (PTC) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.28	-	-
rsp-2(Δ)	-	-	-
rsp-4(Δ)	1.06 ± 0.09	0.26	0.77
rsp-5(Δ)	1.39 ± 0.13	0.09	0.27
rsp-2(Δ) rsp-4(Δ)	-	-	-
rsp-2(Δ) rsp-5(Δ)	-	-	-
rsp-4(Δ) rsp-5(Δ)	1.54 ± 0.30	0.09	0.26
rsp(+)	1 ± 0.5	-	-
rsp-6(Δ)	1.13 ± 0.20	0.7	1
rsp-4(Δ);rsp-6(Δ)	0.94 ± 0.04	0.038	1
rsp-5(Δ);rsp-6(Δ)	1.12 ± 0.76	0.84	1
rsp(+)	1 ± 0.20	-	-
rsp-2(rEx188)	-	-	-
rsp-4(rEx186)	1.20 ± 0.08	0.19	0.74
rsp-4(rEx191)	1.29 ± 0.28	0.23	0.9
rsp-6(rEx187)	1.16 ± 0.13	0.32	1
rsp-6(rEx190)	1.01 ± 0.17	0.94	1

rsp(+)

rsp-1

Strain	Rel. Exp. (+) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.56	-	-
rsp-2(Δ)	0.74 ± 0.40	0.55	1
rsp-4(Δ)	1.20 ± 0.50	0.67	1
rsp-5(Δ)	0.96 ± 0.40	0.92	1
rsp-2(Δ) rsp-4(Δ)	1.49 ± 0.02	0.2	1
rsp-2(Δ) rsp-5(Δ)	1.21 ± 0.50	0.65	1
rsp-4(Δ) rsp-5(Δ)	1.41 ± 0.52	0.41	1
rsp(+)	1 ± 0.25	-	-
rsp-6(Δ)	0.49 ± 0.07	0.027	0.08
rsp-4(Δ);rsp-6(Δ)	0.46 ± 0.11	0.066	0.2
rsp-5(Δ);rsp-6(Δ)	0.22 ± 0.13	0.029	0.09
rsp(+)	1 ± 0.38	-	-
rsp-2(rEx188)	0.56 ± 0.06	0.12	0.36
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	1.34 ± 0.46	0.38	1
rsp-6(rEx190)	1.31 ± 0.51	0.44	1

rsp-2

Strain	Rel. Exp. (+) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.20	-	-
rsp-2(Δ)	-	-	-
rsp-4(Δ)	0.91 ± 0.09	0.5	1
rsp-5(Δ)	1.23 ± 0.21	0.25	0.75
rsp-2(Δ) rsp-4(Δ)	-	-	-
rsp-2(Δ) rsp-5(Δ)	-	-	-
rsp-4(Δ) rsp-5(Δ)	1.13 ± 0.34	0.6	1
rsp(+)	1 ± 0.08	-	-
rsp-6(Δ)	1.14 ± 0.13	0.18	0.55
rsp-4(Δ);rsp-6(Δ)	0.84 ± 0.19	0.37	1
rsp-5(Δ);rsp-6(Δ)	1.05 ± 0.41		
rsp(+)	1 ± 0.43	-	-
rsp-2(rEx188)	-	-	-
rsp-4(rEx186)	0.61 ± 0.08	0.2	0.8
rsp-4(rEx191)	0.80 ± 0.15	0.76	1
rsp-6(rEx187)	1.37 ± 0.28	0.28	1
rsp-6(rEx190)	1.17 ± 0.44	0.66	1

Table 3.2 Continued.

rsp(PTC)

rsp-4

Strain	Rel. Exp. (PTC) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.55	-	-
rsp-2(Δ)	1.04 ± 0.09	0.9	1
rsp-4(Δ)	-	-	-
rsp-5(Δ)	1.35 ± 0.09	0.34	1
rsp-2(Δ) rsp-4(Δ)	-	-	-
rsp-2(Δ) rsp-5(Δ)	1.24 ± 0.06	0.5	1
rsp-4(Δ) rsp-5(Δ)	-	-	-
rsp(+)	1 ± 0.05	-	-
rsp-6(Δ)	1.18 ± 0.19	0.2	0.4
rsp-4(Δ);rsp-6(Δ)	-	-	-
rsp-5(Δ);rsp-6(Δ)	1.48 ± 0.07	0.0025	0.005
rsp(+)	1 ± 0.13	-	-
rsp-2(rEx188)	1.11 ± 0.05	0.25	0.74
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	1.35 ± 0.07	0.015	0.046
rsp-6(rEx190)	1.11 ± 0.31	0.6	1

rsp-5

Strain	Rel. Exp. (PTC) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.57	-	-
rsp-2(Δ)	0.46 ± 0.2	0.19	0.58
rsp-4(Δ)	1.10 ± 0.09	0.79	1
rsp-5(Δ)	-	-	-
rsp-2(Δ) rsp-4(Δ)	0.42 ± 0.17	0.17	0.49
rsp-2(Δ) rsp-5(Δ)	-	-	-
rsp-4(Δ) rsp-5(Δ)	-	-	-
rsp(+)	1 ± 0.36	-	-
rsp-6(Δ)	0.69 ± 0.17	0.25	0.5
rsp-4(Δ);rsp-6(Δ)	0.36 ± 0.17	0.38	0.75
rsp-5(Δ);rsp-6(Δ)	-	-	-
rsp(+)	1 ± 0.23	-	-
rsp-2(rEx188)	1.26 ± 0.11	0.15	1
rsp-4(rEx186)	1.9 ± 0.44	0.035	0.18
rsp-4(rEx191)	2.65 ± 0.78	0.033	0.17
rsp-6(rEx187)	1.23 ± 0.09	0.28	1
rsp-6(rEx190)	1.04 ± 0.50	0.91	1

rsp(+)

rsp-4

Strain	Rel. Exp. (+) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.51	-	-
rsp-2(Δ)	0.89 ± 0.47	0.8	1
rsp-4(Δ)	-	-	-
rsp-5(Δ)	1.35 ± 0.68	0.27	0.8
rsp-2(Δ) rsp-4(Δ)	-	-	-
rsp-2(Δ) rsp-5(Δ)	1.48 ± 0.5	0.31	0.94
rsp-4(Δ) rsp-5(Δ)	-	-	-
rsp(+)	1 ± 0.17	-	-
rsp-6(Δ)	0.97 ± 0.04	0.75	1
rsp-4(Δ);rsp-6(Δ)	-	-	-
rsp-5(Δ);rsp-6(Δ)	0.79 ± 0.2	0.29	0.58
rsp(+)	1 ± 0.23	-	-
rsp-2(rEx188)	0.81 ± 0.06	0.24	0.73
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	1.83 ± 0.88	0.19	0.57
rsp-6(rEx190)	1.44 ± 0.45	0.2	0.61

rsp-5

Strain	Rel. Exp. (+) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.56	-	-
rsp-2(Δ)	1.88 ± 1.05	0.27	0.81
rsp-4(Δ)	2.17 ± 0.60	0.07	0.21
rsp-5(Δ)	-	-	-
rsp-2(Δ) rsp-4(Δ)	3.4 ± 0.12	0.002	0.006
rsp-2(Δ) rsp-5(Δ)	-	-	-
rsp-4(Δ) rsp-5(Δ)	-	-	-
rsp(+)	1 ± 0.18	-	-
rsp-6(Δ)	0.96 ± 0.17	0.8	1
rsp-4(Δ);rsp-6(Δ)	0.89 ± 0.39	0.37	0.75
rsp-5(Δ);rsp-6(Δ)	-	-	-
rsp(+)	1 ± 0.37	-	-
rsp-2(rEx188)	0.60 ± 0.05	0.14	0.68
rsp-4(rEx186)	0.51 ± 0.12	0.096	0.48
rsp-4(rEx191)	0.85 ± 0.12	0.64	1
rsp-6(rEx187)	1.25 ± 0.19	0.44	1
rsp-6(rEx190)	1.41 ± 0.46	0.3	1

rsp(+)

Table 3.2 Continued.

rsp(PTC)

rsp-6

Strain	Rel. Exp. (PTC) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.62	-	-
rsp-2(Δ)	1.26 ± 0.23	0.53	1
rsp-4(Δ)	1.16 ± 0.05	0.68	1
rsp-5(Δ)	1.44 ± 0.25	0.32	1
rsp-2(Δ) rsp-4(Δ)	1.33 ± 0.06	0.42	1
rsp-2(Δ) rsp-5(Δ)	1.43 ± 0.29	0.34	1
rsp-4(Δ) rsp-5(Δ)	1.37 ± 0.19	0.39	1
rsp(+)	-	-	-
rsp-6(Δ)	-	-	-
rsp-4(Δ);rsp-6(Δ)	-	-	-
rsp-5(Δ);rsp-6(Δ)	-	-	-
rsp(+)	1 ± 0.39	-	-
rsp-2(rEx188)	0.55 ± 0.06	0.12	0.12
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	-	-	-
rsp-6(rEx190)	-	-	-

rsp-7

Strain	Rel. Exp. (PTC) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.81	-	-
rsp-2(Δ)	0.98 ± 0.10	0.97	1
rsp-4(Δ)	0.68 ± 0.05	0.54	1
rsp-5(Δ)	1.31 ± 0.06	0.55	1
rsp-2(Δ) rsp-4(Δ)	0.65 ± 0.10	0.5	1
rsp-2(Δ) rsp-5(Δ)	1.21 ± 0.06	0.67	1
rsp-4(Δ) rsp-5(Δ)	0.68 ± 0.09	0.54	1
rsp(+)	1 ± 0.06	-	-
rsp-6(Δ)	0.65 ± 0.17	0.03	0.09
rsp-4(Δ);rsp-6(Δ)	0.35 ± 0.02	0.0007	0.002
rsp-5(Δ);rsp-6(Δ)	1.01 ± 0.65	0.97	1
rsp(+)	1 ± 0.09	-	-
rsp-2(rEx188)	0.61 ± 0.03	0.002	0.012
rsp-4(rEx186)	0.81 ± 0.02	0.025	0.13
rsp-4(rEx191)	1.24 ± 0.42	0.38	1
rsp-6(rEx187)	0.90 ± 0.23	0.64	1
rsp-6(rEx190)	0.78 ± 0.23	0.19	0.97

rsp(+)

rsp-6

Strain	Rel. Exp. (+) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.63	-	-
rsp-2(Δ)	0.69 ± 0.52	0.55	1
rsp-4(Δ)	1.44 ± 0.12	0.3	1
rsp-5(Δ)	1.54 ± 0.44	0.29	1
rsp-2(Δ) rsp-4(Δ)	1.41 ± 0.21	0.42	1
rsp-2(Δ) rsp-5(Δ)	1.16 ± 0.37	0.73	1
rsp-4(Δ) rsp-5(Δ)	2.08 ± 1.09	0.21	1
rsp(+)	-	-	-
rsp-6(Δ)	-	-	-
rsp-4(Δ);rsp-6(Δ)	-	-	-
rsp-5(Δ);rsp-6(Δ)	-	-	-
rsp(+)	1 ± 0.43	-	-
rsp-2(rEx188)	0.72 ± 0.05	0.33	0.33
rsp-4(rEx186)	-	-	-
rsp-4(rEx191)	-	-	-
rsp-6(rEx187)	-	-	-
rsp-6(rEx190)	-	-	-

rsp-7

Strain	Rel. Exp. (+) ± SD	P-value	Corrected P-value
rsp(+)	1 ± 0.66	-	-
rsp-2(Δ)	0.76 ± 0.39	0.62	1
rsp-4(Δ)	1.97 ± 0.18	0.07	0.42
rsp-5(Δ)	1.43 ± 0.11	0.55	1
rsp-2(Δ) rsp-4(Δ)	1.90 ± 0.06	0.08	0.47
rsp-2(Δ) rsp-5(Δ)	0.98 0.31	0.97	1
rsp-4(Δ) rsp-5(Δ)	1.40 ± 0.28	0.39	1
rsp(+)	1 ± 0.11	-	-
rsp-6(Δ)	1.02 ± 0.14	0.88	1
rsp-4(Δ);rsp-6(Δ)	1.06 ± 0.15	0.65	1
rsp-5(Δ);rsp-6(Δ)	0.77 ± 0.25	0.24	0.72
rsp(+)	1 ± 0.32	-	-
rsp-2(rEx188)	0.53 ± 0.05	0.07	0.33
rsp-4(rEx186)	0.39 ± 0.04	0.03	0.16
rsp-4(rEx191)	0.52 ± 0.04	0.14	0.71
rsp-6(rEx187)	0.39 ± 0.04	0.68	1
rsp-6(rEx190)	0.52 ± 0.04	0.63	1

Chapter 4: Conclusions and Future Directions

The similar number of genes and the similar gene functions in dramatically different eukaryotes suggests eukaryotes regulate patterns of gene expression in different ways. Such regulation is important to achieve the striking diversity of species observed within the animal and plant kingdoms. One mechanism to regulate gene expression is alternative splicing. Alternative splicing is now appreciated as a far more common event than first hypothesized. Regulation of alternative splicing is therefore of great importance; uncovering how regulators of alternative splicing are themselves regulated furthers our understanding of gene expression and its effects.

mRNAs encoding splicing regulators are themselves often alternatively spliced. These observations suggest alternative splicing is a means to regulate proteins that regulate splicing. Intriguingly, one of the common outcomes of alternative splicing is the generation of mature mRNAs that are rapidly destroyed. Introduction of premature-termination codons (PTCs) in alternatively spliced mature mRNAs subjects PTC-containing mRNAs to the nonsense-mediated mRNA decay pathway (NMD).

I observed the *rsp* family of mRNAs in *C. elegans* is alternatively spliced to introduce PTCs into its mature mRNA transcripts. *Rsp* mRNAs encode proteins orthologous to the well-known mammalian SR protein family of splicing regulators. This thesis describes my observations of *rsp* alternative splicing patterns and experiments designed to address why and how these patterns of splicing are regulated. Presumably, regulation of splicing factors, including RSP proteins, affects downstream targets. Therefore regulation of splicing regulators is bound to have profound consequences on gene expression in eukaryotes. In the

next section, I will briefly describe my major conclusions regarding *rsp* splicing regulation and then propose experiments to address lingering questions regarding *rsp* splicing regulation and function.

Alternative splicing of rsp mRNAs

I observed splicing of *rsp* mRNAs that generate PTCs and are degraded by NMD. *SR* mRNAs are alternatively spliced throughout eukaryotes, including alternative splicing to generate mature *SR* mRNAs containing PTCs. These alternative splicing events are now thought as means to regulate gene expression, and as demonstrated by the number of organisms these events are observed in, appear highly conserved.

Developmental regulation of rsp mRNA and splicing patterns

I observed regulation of *rsp* mRNAs throughout *C. elegans* development. I draw two major conclusions: first, my observations suggest splicing regulators are needed at different levels throughout development, and two, regulation of *rsp* splicing appears to be a means to regulate *rsp* expression throughout development. I hypothesize changes in alternative splicing are a means to affect small changes in functional *rsp* mRNAs translated into functional proteins. What controls the large changes in *rsp* mRNAs throughout development is unclear.

Developmental regulation of* RSP-6 *protein abundance

Changes in *rsp* mRNA abundance throughout development suggest possible corresponding changes in protein abundance. I sought to test RSP protein abundance by generating antibodies to measure RSP protein abundance throughout development. My results suggest at least one RSP protein, RSP-6, is developmentally regulated. I observed decreases in the abundance of RSP-6 protein subsequent to decreases in corresponding *rsp-6* mRNA. Though decreases in RSP-6 happen long after decreases in *rsp-6* mRNA, it is possible the downregulation of both mRNA and protein are related. RSP-6 protein may be stable or unknown factors may stabilize RSP-6 long after decreases in *rsp-6(+)* mRNA.

Even if there is no relation between *rsp-6(+)* mRNA and RSP-6 protein, I observed significant developmentally regulated changes in RSP-6 protein abundance. Knowing the roles of RSP orthologues, it is an interesting, though not unexpected observation that a splicing factor is developmentally regulated. The orthologue of RSP-6, SRSF3, is involved in both splicing and interacts with RNA Pol II during transcription [73]. Thus, *C. elegans* RSP-6 protein may be involved in several processes including regulating splicing; regulation of these processes throughout development could have profound consequences on downstream gene expression.

RUST of rsp mRNAs

My initial observations of *rsp* mRNAs suggested their splicing patterns are likely regulated. Prior observations of similar splicing patterns, such as *C. elegans rpl* alternative splicing, suggested the regulation of *rsp* alternative splicing might be a means to maintain

protein homeostasis through a mechanism described as RUST. My observations are consistent with the RUST model. I performed experiments to measure the splicing patterns of *rsp* mRNAs in mutants that modulate RSP expression. I observed splicing patterns consistent with a model where changes in splicing patterns are correlated with maintaining protein homeostasis. Furthermore, I observed several RSP proteins appear to affect the splicing patterns of other *rsp* pre-mRNAs.

Future Directions

What explains the decrease in rsp mRNA abundance between the embryo and L1 transition?

The relative abundance of *rsp* mRNAs decreases dramatically early in *C. elegans* development. I hypothesize the corresponding change in the splicing patterns of four of the *rsp* mRNAs is a means to further decrease the abundance of *rsp(+)* mRNAs. The decrease in *rsp(total)* mRNA caused by shifting splicing patterns to increase the proportion of *rsp(PTC)* mRNAs can only explain a fraction of the decrease in *rsp(total)* mRNA. Thus, the changes in mRNA abundance are likely explained by one of the following reasons: 1) a decrease in *rsp* transcription rates, 2) an increase in the rate of *rsp* mRNA degradation, or 3) a combination of decreased transcription and increased degradation rates.

I hypothesize the likeliest explanation for the decrease in total *rsp* mRNA between embryo and L1 is a decrease in transcription rates of *rsp* mRNAs. I initially attempted to assay *rsp* transcription rates by performing nuclear run on experiments in both embryo and L1-stage samples. The Kennedy lab developed a robust protocol to measure transcription

rates in *C. elegans* embryos [179], but I was unsuccessful in adapting this technique to also measure transcription rates in L1s. Though I was initially unsuccessful, it is possible this technique might work to measure L1 transcription rates given enough time to optimize sample preparation and other issues I encountered. Other possibilities include less direct approaches of measuring the occupancy of RNA Polymerase II (Pol II) along the length of *rsp* loci by Pol II chromatin immunoprecipitation (ChIP). Enrichments of Pol II near alternatively spliced exons are known, and it is possible I might observe similar enrichments at *rsp* alternative exons [52]. Enrichments at exons are generally thought as stalling or pausing of Pol II, and have been interpreted as decreases in transcription rates [52]. Pol II enrichments at *rsp* loci would fit with the current understanding of connections between transcription rates and alternative splicing.

Increasing the degradation rate of *rsp* mRNAs is another possible reason for changes in *rsp* mRNA abundance. There are a number of pathways that could ultimately degrade unneeded *rsp* mRNAs, including pathways that generate small RNAs that might be directed against *rsp* mRNAs. I could use mutants for many degradation pathways and ask if *rsp(total)* mRNAs are stabilized in the embryo to L1 transition compared to wild type.

It is also possible the relatively large amount of *rsp* mRNA in embryos is deposited maternally. Such mRNAs may be stabilized until the onset of zygotic transcription. After such time they may no longer be needed and could be immediately degraded. The abundance of *rsp* mRNA transcribed in zygotes is most likely far lower than that which would be maternally deposited, thus this would explain the observed decreases in *rsp* mRNAs between embryo and L1.

What causes splicing patterns to change in C. elegans development?

I hypothesize changes in the splicing patterns of *rsp* and *hrpf-1* mRNAs are mechanisms worms use to fine-tune the level of functional *rsp* and *hrpf-1* mRNAs. Excess pre-mRNA spliced to include a PTC is rapidly destroyed by NMD and therefore unneeded mRNA is shuttled to this pathway. What causes this change in *rsp* and *hrpf-1* mRNA splicing patterns is unknown. Many of the *rsp* splicing patterns appear governed by RUST, and such a mechanism might explain part of the developmental regulation of *rsp* mRNAs I observed. Since I observed several *rsp* splicing patterns regulated by other RSP proteins, there may be a combination of feedback of one RSP regulating its own splicing patterns, but also the splicing patterns of other *rsp* mRNAs. Therefore RSP proteins may regulate their own and each other's splicing patterns throughout development.

The change in *rsp(total)* mRNA between embryo and L1 could be due to decreases in transcription rates. Such decreases in transcription might also explain the changes in the splicing patterns of *rsp* mRNAs between embryo and L1. The kinetic model of alternative splicing suggests slower transcription rates allow suboptimal splice sites to be more easily recognized, increasing the rate of alternative splicing. If Pol II is slowed along *rsp* loci between the embryo and L1 transition, this might allow both less *rsp* pre-mRNA to be generated, as well as potentially influencing the splicing patterns of *rsp* mRNAs. Nuclear run on assays or Pol II ChIP experiments might demonstrate a link between *rsp* transcription rates and splicing.

Changes in *rsp* splicing patterns throughout development might also be influenced by chromatin modifications. Recent data in other organisms suggest DNA modifications, along

with changes in the abundance of splicing regulators, can have profound implications on regulating tissue-specific alternative splicing (reviewed in [13]). It is possible chromatin modifications change throughout the course of worm development. Recent experiments suggest changes in chromatin modifications, particularly H3K36me3 states [94], may have effects on splicing patterns of specific mRNAs. I could assay H3K36me3 and other chromatin states associated with splicing along *rsp* loci at different developmental stages by chromatin IP (ChIP) with antibodies that recognize specific histone marks and perform qRT-PCR to measure association of histone marks at *rsp* loci.

Are all rsp splicing patterns affected by starvation? Do other physiological stresses affect rsp mRNAs?

Splicing patterns of *rsp-6, rsp-7,* and *hrpf-1* change under starvation conditions [52,53]. In starved L1-stage worms, the PTC isoform proportion increases for *rsp-6, rsp-7,* and *hrpf-1* mRNAs. Correlated with these changes in splicing is increased Pol II occupancy of introns flanking the alternatively spliced exons of *rsp-6, rsp-7,* and *hrpf-1*. Worms fed subsequent to starvation quickly change their splicing patterns to decrease the *rsp-6, rsp-7,* and *hrpf-1* PTC isoforms, as well as decreasing Pol II occupancy in adjacent introns. These observations suggest during stress, transcription decreases at *rsp* and *hrpf-1* loci and splicing patterns change to further decrease *rsp* and *hrpf-1* mRNAs by shuttling more pre-mRNA to the NMD pathway. In *Arabidopsis,* a number of physiological stresses, including temperature and changes in hormones levels affect the splicing patterns of many *SR* mRNAs [39,46]. These observations suggest one function of *SR* mRNA alternative splicing is to downregulate

SR proteins in times of stress by changing the proportion of mature mRNA encoding functional SR proteins.

I could repeat starvation experiments on the other *rsps* I observed as regulated throughout *C. elegans* development. The observations above suggest a decrease in functional splicing regulators during times of stress. I could therefore also test other stress conditions in *C. elegans* and ask what happens to *rsp* and *hrpf-1* splicing patterns under those conditions, as well as whether Pol II stalls near sites of alternative splicing.

Why is there downregulation of rsp(+) and hrpf-1(+) mRNA between embryo and L1?

No matter the cause of downregulation of *rsp* mRNA between embryo and L1, my results suggest there is less need for *rsp(+)* mRNAs in L1 and most of the remainder of the *C. elegans* lifespan. Despite differences in the splicing patterns of *hrpf-1* and *rsp* mRNAs, there is a similar decrease in the proportion of *rsp(+)* and *hrpf-1(+)* between the embryo and L1 transition. Why do *rsp* and *hrpf-1* splicing regulators, which in general work in different ways, increase proportions of their mRNA pools that are nonfunctional in L1?

A drastic decrease in *rsp(+)* mRNA in L1 worms may cause decreases in the abundance of RSP proteins. If so, this might have profound impacts on global patterns of alternative splicing. Changes in alternative splicing may be a crucial aspect of how worms properly develop. RSP proteins may set up pathways of splicing early in development that are either self-enforcing or trigger cascades of alternative splicing of downstream genes. Such patterns are observed in *Drosophila* in experiments that dissected the fly sex-determination pathway [168,169]. My experiments also suggest a second possibility that RSP

proteins may have long half-lives. Therefore a change in RSP protein level is not required until later in development, but because of a long protein half-life, mRNA levels are turned down in advance.

What are the targets of RSP proteins in C. elegans?

Knowing what downstream targets RSP proteins regulate is a crucial way to understand why there are changes in RSP protein levels. The simplest approach to identify RSP protein targets would be to run RNA-seq analysis in *rsp* mutant strains and determine which mRNAs change in abundance. A more informative approach would be to run RIP-seq (RNA-immunoprecipitation) experiments to identify direct targets of RSP proteins. The RSP-6 antibody used in this thesis is unlikely to reproducibly precipitate endogenous RSP-6 protein (despite extensive optimization, I was unable to show RSP-6 is precipitated with anti-RSP-6). A different approach would be to generate FLAG or other-tagged transgenic strains and assess what mRNAs these tagged-proteins associate with. A recent report of SR associated mRNAs in mammals could be used to cross-reference and determine if there are common themes in worm and mammalian SR target mRNAs [50].

Because I observed RSP-6 protein abundance is developmentally regulated, it is possible RSP-6 associates with and affects the splicing of target mRNAs in different ways throughout development. A large fraction of alternatively spliced *C. elegans* mRNAs change isoform abundance throughout development [8]. Therefore it may be most informative to perform RIP-seq at each stage, or at least the earliest stages, of worm development.

How are rsp mRNAs regulated by RUST?

The simplest explanation of RUST is one where proteins direct the splicing patterns of their own pre-mRNAs. A number of splicing factors are known to regulate their own mRNAs by such a mechanism. Mammalian proteins SRSF3 and SRSF4 (orthologues of RSP-6 and RSP-1, respectively) associate with their own mRNAs. These mRNAs are alternatively spliced to generate PTC-containing mature mRNAs similar to the *rsps*. Therefore it is possible a similar mechanism exists to regulate *C. elegans rsp* mRNAs.

I was unable to immunoprecipitate endogenous *C. elegans* RSP-6 or transgenic GFP::RSP proteins using anti-RSP-6 or anti-GFP antibodies, respectively. Other epitope tags could be used to test if RSP proteins associate with their own and each other's pre-RNAs. These tagged strains could also be used in the RIP-seq experiments proposed earlier.

If I was unable to pull down endogenous or tagged RSP proteins by the above methods, I could assess whether RSP proteins associate with *rsp* mRNAs *in vitro*. In the course of generating antibodies for RSP proteins, I made a series of GST, HIS, and Maltose-tagged constructs for RSP-2 and RSP-6, including full-length proteins. Generation of other tagged RSP proteins is not technically difficult, and by generating recombinant RSP proteins I could assess their ability to interact with labeled *rsp* mRNAs *in vitro*. It is likely RSP proteins associate with their own mRNAs, and through *in vitro* approaches I might be able to identify where such interactions occur. Similar experiments showed recombinant human hnRNP L protein associates with exon 6A of *hnRNP L* mRNA in experiments showing *hnRNP L* alternative splicing is regulated through RUST [162].

Do RSP proteins associate with each other's pre-mRNAs?

Assuming the experiments described above worked, it would be trivial to extend those methods to assess if RSP proteins also interact with each other's mRNAs. Immunoprecipitation followed by RT-PCR would allow me to determine if one or more RSP protein interacts with other mRNAs. Recent experiments show SRSF3 (mammalian orthologue of RSP-6) interacts with two other *SRSF* mRNAs [50]. These results suggest I might observe similar results in worms.

Concluding Remarks

Alternative splicing regulation is clearly an important factor in proper gene expression and is implicated in many developmental and tissue-specific processes. This thesis demonstrates several ways I observed a family of splicing regulators are themselves regulated through alternative splicing. The patterns I observed indicate there may be many ways these events are regulated and I propose future experiments that could further our understanding of the mechanism by which these splicing events work.

Materials and Methods

Strains

Strains used were N2 (wild type), TR1335 [*smg-5(r860)*], TR1436 [*smg-2(r908)*], TR1438 [*smg-1(r910)*], TR1953 [*smg-3(r930)*], TR2482 [*mIn1[mIs14],dpy-10(e128);smg-2(r908)*], TR2483 [*nT1[qIs51]; smg-2(r908)*], TR2485 [*nT1[qIs51]/rsp-6(ok798);smg-2(r908)*], TR2488 [*mIn1[mIs14], dpy-10(e128)/rsp-4(tm837);smg-2(r908)*], TR2489 [*rsp-4(tm837);smg-2(r908)*] TR2498 [*rsp-2(tm952);smg-2(r908)*], TR2499 [*rsp-5(ok324);smg-2(r908)*], TR2501 [*mIn1[mIs14], dpy-10(e128)/rsp-2(tm952);smg-2(r908)*], TR2502 [*mIn1[mIs14],dpy-10(e128)/rsp-5(ok324);smg-2(r908)*], TR2556 [*smg-2(r908);gfp::rsp-4(rEx186)*], TR2557 [*smg-2(r908);gfp::rsp-6(rEx187)*] TR2558 [*smg-2(r908);gfp::rsp-6(rEx190)*], TR2562 [*smg-2(r908);gfp::rsp-2(rEx188)*], TR2599 [*rsp-5(ok324) rsp-2(tm952);smg-2(r908)*], TR2601 [*rsp-4(tm837) rsp-2(tm952);smg-2(r908)*], TR2653 [*rsp-4(tm837) rsp-5(ok324);smg-2(r908)*], TR2654 [*rsp-5(ok324);nT1(qIs41)/rsp-6(ok798);smg-2(r908)*], TR2656 [*rsp-4(tm837);nT1[qIs51]/rsp-6(ok798);smg-2(r908)*], and TR2661 [*smg-2(r908);gfp::rsp-4(rEx191)*].

C. elegans strains were raised at 20°C. Strains for developmental analyses were grown in liquid cultures by standard methods and hypochlorite treated to isolate embryos, synchronized, then grown to gravid adult stage and hypochlorite treated again to isolate embryos. Embryos were collected by sucrose flotation and a portion was used as a sample for embryo stage. The remaining portion was grown to appropriate stages and collected by sucrose flotation. Strains for autoregulation feedback analysis were grown on NGM plates and handpicked (~300 L4 worms per sample). For cross-regulation analysis, *rsp-6(ok798)*

containing strains, transgenic strains, and their controls were grown on NGM plates and handpicked as above. All remaining strains for cross-regulation analysis were grown in liquid cultures and hypochlorite treated as above, synchronized, grown to L4-stage and sucrose floated.

DNA Cloning

Plasmids containing *rsp-2* (TR#493), *rsp-4* (TR#494), and *rsp-6* (TR#492) were generated using full-length genomic sequences for each respective *rsp*, except the exons flanking the PTC isoform were fused together, removing the PTC region, and cloned into TR#429 (described in [159]) to express GFP::RSP fusions driven by the *eft-3* promoter and tagged with *unc-54* 3' UTR. Specifically, starting with their start codons, TR#493 expresses nucleotides 1-430 and 720-1283 of *rsp-2*, TR#494 expresses nucleotides 1-400 and 878-1167 of *rsp-4*, and TR#492 expresses nucleotides 1-392 and 832-1125 of *rsp-6*.

Plasmids for RNA probes used in northern blots were generated from cDNA cloned into pGEM T-Easy (Promega) vector (except TR#509 was cloned into pGEM T vector) and correspond to the following genomic DNA nucleotide numbers/sequences: *rsp-1* (TR#522) nucleotides 461-533 and 1045-1118, *rsp-2* (TR#488) nucleotides 1-120 and 173-430, *rsp-4* (TR#489) nucleotides 1-206 and 263-400, *rsp-5* (TR#490) nucleotides 1-188 and 240-491, *rsp-6* (TR#491) nucleotides 1-54 and 108-238, and *eft-3* (TR#509) nucleotides 27-518 (*eft-3* plasmid cloned by Amy Hubert).

TR#491 was cloned into pGEX 4T-1 to generate GST-tagged RSP-6 (TR#527) to generate anti-RSP-6 sera. Serum containing anti-RSP-6 antibodies was purified against full-

length *rsp-6* cDNA cloned into pHMTc (to generate His and Maltose tagged RSP-6, TR#528). pHMTc is a derivative of the pMal-c2X vector (New England Biolabs) obtained from the Wickens lab, see [180] for reference.

Nematode Transformation

TR2556, TR2557, TR2558, TR2562, and TR2661 were generated using plasmids TR#s 492-494 by standard microinjection methods at 5 or 10 ng/µl, plus pRF4[*rol-6(su1006)*] used as a marker plasmid at 50ng/µl, and a 1 kb DNA ladder (New England Biolabs) to increase the complexity of extrachromosomal arrays at either 45 or 50 ng/µl (for a total DNA concentration of 100ng/µl per injection).

RNA Preparation and cDNA synthesis

Worm samples were frozen in Trizol (Invitrogen) at -80°C until processing. RNA samples were processed with 3-5 cycles of heating samples in Trizol to 65°C and flash-freezing in liquid Nitrogen, and extracting RNA according to manufacturer's instructions. For RT-PCR and qRT-PCR, samples were quantified by UV absorbance at 260 nm and diluted to 1 µg total RNA for first-stand synthesis using Super Script III (Invitrogen) reverse transcriptase and random hexamers.

RT-PCR and qRT-PCR

cDNA samples for RT-PCR were amplified with GoTaq polymerase (Promega) using primer pairs specific for mature mRNA that amplified both the PTC and + transcripts. For qRT-

PCR, cDNA samples were diluted 5-fold after first strand synthesis and amplified with Power SYBR Green (Ambion) on a Bio-Rad IQ5 Cycler. Primer pairs were designed to amplify only one mature *rsp* isoform (PTC, +, or total) and cross exon-exon boundaries.

Northern Blotting

RNA samples were purified as above and further processed with Qiagen RNeasy columns prior to quantification by UV absorbance at 260 nm. 8-10 μg total RNA were resuspended in a glyoxal/DMSO loading mix (described in [181]), separated on 3% Agarose gels run for approximately 16-20 hours in 1x BPTE buffer, and blotted to Zeta-Probe membranes (Bio-Rad) cross-linked with a stratalinker. Antisense RNA probes were generated using linearized plasmid templates described above with a Maxiscript kit (Ambion) and α-P^{32} incorporation. Free nucleotides were removed from probes with G-50 microspin columns (GE Healthcare). Blots were probed overnight at 68°C in Ultrahyb Hybridization Buffer (Ambion) and washed as described in the Maxiscript kit booklet protocol (Ambion). Blots were exposed overnight to a storage phosphor screen, scanned by a Storm scanner, and quantified using ImageQuant software. The abundance of *rsp* transcripts were normalized to *eft-3* mRNA.

Antibodies

Anti-RSP-6 serum was generated by immunizing rabbits with a recombinant glutathione *S*-transferase-tagged RSP-6 (amino acids 1-61 inserted into pGEX 4T-1) protein purified with glutathione-sepharose 4B (GE Healthcare). Serum containing RSP-6 antibodies (UWM248) was purified against a maltose-tagged full-length RSP-6 recombinant protein (amino acids 1-

179 cloned into pHMTc) purified with amylose resin (New England Biolabs) bound to CNBr-activated sepharose 4 fast flow beads (GE Healthcare). Purified RSP-6 antibodies were desalted and concentrated (Amicon Ultra Centrifugal Filter 10k MWCO).

Western Blotting

Worm protein samples for developmental analysis were washed several times in M9 buffer and resuspended in 1x PBS, sonicated, quantified (Bio-Rad Protein Assay), and 15 μg was added to equal volumes of Laemelli Sample Buffer (Bio-Rad) plus 5% β-mercaptoethanol. Protein samples for feedback regulation analysis were handpicked and approximately 300 L4 worms were picked into M9 buffer and processed as above. Samples were heated to 95°C for 10 minutes before separation on 12% SDS-PAGE gels and transferred to Immobilon-FL PVDF (Millipore) using a Trans-Blot semidry transfer cell (Bio-Rad). Blots were blocked and probed in a 1:1 mixture of Odyssey blocking buffer (LiCor) and 1x PBS. Blots were probed with anti-RSP-6 (1:500; incubated overnight at 4°C), anti-GFP (1:2500; clone JL-8 Clonetech), or anti-actin (1:4000; clone C4 MD Biomedicals) primary antibodies in blocking buffer plus 0.1% Tween. Blots were incubated in secondary rabbit IgG (1:10,000; LiCor) IRDye 680 and mouse IgG (1:20,000; LiCor) IRDye 800 CW antibodies in blocking buffer plus 0.1% Tween and 0.01% SDS. Blots were scanned using a LiCor Odyssey scanner and quantified with manufacturer's software.

Table M.1. RT-PCR primers that amplify *rsp* mature mRNA transcripts.

PCR Reaction	Forward Primer	Forward Primer Sequence	Reverse Primer	Reverse Primer Sequence
rsp-1	rsp-1 e2/e3 for	ACCACCAAGAAGAGAAAGTCG	rsp-1 e4/e5 rev	GTTGCGAAGCAGAGAAGTGC
rsp-2	rsp-2 dT left	ACCTCTAAACATGGTTCGTGTCTACATC	rsp-2 e3/e4 rev	CATACGATTGCTTGGTTGAC
rsp-3 A	rsp-3 e1 for 1	GTCCGAGAGAAGGAAGTCG	rsp-3 e2 rev 1	GCCACATCCGCATAGCAGAC
rsp-3 B	rsp-3 e1 for 1	GTCCGAGAGAAGGAAGTCG	rsp-3 e3 rev 1	GAACGAGAACGACGAGGTGAG
rsp-3 C	rsp-3 e2 for 1	GTCTGCTATGCGGATGTGGC	rsp-3 e3 rev 1	GAACGAGAACGACGAGGTGAG
rsp-4	r4 FOR 1	TTTTATGAACGTCGTGATGCTG	r4 REV 1	TAGCGGGAGTTGGAACGG
rsp-5	r5 e1b for (1)	GCTGAAGACGCATGCCATGAT	r5 e4 a/b rev	ATCCACTTCGGGTTCTGCTTCG
rsp-6	rsp-6 e2/e3 for	GACGGATCAAGAATCTGTGGTGTC	rsp-6 e4 rev	GACTGCGTTCACGACTGCG
rsp-7	rsp-7 e3 for	CAACAGTGGTGGATCTACAAC	rsp-7 e4/e5 rev	TTCTGGCGTACATTACCTCTC

Table M.3. Genotyping primers for *rsp* deletion alleles.

rsp-2(tm952)	**Forward Primer**	**tm952 for**
	Forward Primer Sequence	TCAACCTGAACACCTACGCA
	Reverse Primer	**tm952 rev wt**
	Reverse Primer Sequence	TACCACGAAACCAAATCCG
	Reverse Primer 2	**tm952 rev mut**
	Reverse Primer 2 Sequence	CCCACTGAGTTGCTGCTGTAG
rsp-4(tm837)	**Forward Primer**	**tm837 for**
	Forward Primer Sequence	CTGGAGGAATGGTGGATGC
	Reverse Primer	**tm837 rev wt**
	Reverse Primer Sequence	CAGCGAAGTTAGACCATTGATA
	Reverse Primer 2	**tm837 rev mut**
	Reverse Primer 2 Sequence	CACTGTCGCCGCTGGTT
rsp-5(ok324)	**Forward Primer**	**ok324 for**
	Forward Primer Sequence	CCAGTCATTTTGAGCACATTCG
	Reverse Primer	**ok324 rev wt**
	Reverse Primer Sequence	GACAAGGCGCATAGAGCTGAA
	Reverse Primer 2	**ok324 rev mut**
	Reverse Primer 2 Sequence	TCAAGGAAACATCACTCGCAGG
rsp-6(ok798)	**Forward Primer**	**ok798 for**
	Forward Primer Sequence	ACTTTGACTGGTTTTGCTATTTC
	Reverse Primer	**ok798 rev wt**
	Reverse Primer Sequence	GCATCACTCGGCAGACC
	Reverse Primer 2	**ok798 rev mut**
	Reverse Primer 2 Sequence	CCAAAACTCTTACCTTGAACGA

Table M.3. qRT-PCR primers for *rsp* mRNA quantification.

PCR Reaction	Forward Primer	Forward Primer Sequence	Reverse Primer	Reverse Primer Sequence
rsp-1(PTCshort)	rsp-1 PTC(+) short for 1	CAAGGATTGCATATGGGATCTC	rsp-1 e4/e5 rev	GTTGCGAAGCAGAGAAGTGC
rsp-1(PTClong)	rsp-1 PTC(+) long for 1	CACGCACATTTTTCACAACACC	rsp-1 e4/e5 rev	GTTGCGAAGCAGAGAAGTGC
rsp-1(+)	rsp-1 e2/e3 for	ACCACCAAGAAGAGAAAGTCG	rsp-1 PTC(-) rev 1	CTTGGTCTTTGAGATCCTGCCA
rsp-1(total)	rsp-1 e2 for	GGTGGTGGCGGTGGAGG	rsp-1 e2/e3 rev	ATGGACGACCGTAGCGACTTTC
rsp-2(PTC)	rsp-2 PTC(+) For 3	TCGTTACAGTTGGCAGCAGCA	rsp-2 PTC(+) rev 3	GGTATTGATTGTTTGTGAAGTAGGG
rsp-2(+)	rsp-2 PTC(-) for 2	CAGTTGGCAGGACATCAAGGAC	rsp-2 PTC(+/-) rev 2	TGAAGCATACGATTGCTTGGTTG
rsp-2(total)	rsp-2 for mat 2	ACGGATTTGGTTTCGTGGATT	rsp-2 rev mat 2	GAATGACTCTCTCTCCACAAAG
rsp-4(PTC)	rsp-4 PTC(+/-) for 3	CTGAGCACGCATTGGACCG	rsp-4 PTC(+) rev 3	CTGGCTCTGCTGCTGTTCTTC
rsp-4(+)	rsp-4 PTC(+/-) for 3	CTGAGCACGCATTGGACCG	rsp-4 PTC(-) rev 3	CTCGGTGAGCGGGATCTTCT
rsp-4(total)	rsp-4 for mat 1	CGAACTTTTGAGCGATACGG	rsp-4 rev mat 1	GCTCAGCATCACGACGTTC
rsp-5(PTC)	rsp-5 PTC(+) for 4	CGTTCAGTCCCCACAGTTACC	rsp-5 PTC(+) rev 4	CTATCTCCGCCGTTCTTTGG
rsp-5(+)	rsp-5 PTC(-) for 3	CGTTCCCGCTCTCGTGAT	rsp-5 PTC(-) rev 3	CGACGCTTCACGCTTCAA
rsp-5(total)	rsp-5 for mat 2	GTATGGATTTGCCTTTGTCG	rsp-5 rev mat 2	GACAAGGCGCATAGAGCTTC
rsp-6(PTCshort)	rsp-6 PTC(+) for 4	GTCTTCACGACACATTCTTCTCTTC	rsp-6 short rev	TCTCCTCGGTATGGCGATTGA
rsp-6(PTClong)	rsp-6 PTC(+) for 4	GTCTTCACGACACATTCTTCTCTTC	rsp-6 PTC(+) long rev 3	CCTCGGTATGGCGATCTTTCC
rsp-6(+)	rsp-6 PTC(+/-) for 2	GACGGATCAAGAATCTGTGGTG	rsp-6 PTC(-) rev 1	CGGTATGGCGATCTGTCTCTTC
rsp-6(total)	rsp-6 for mat 2	ATGGACGCCAAGGTGTACG	rsp-6 rev mat 2	CGATCAAAAATCTCTTCGAGTTC
rsp-7(PTC)	rsp-7 e3 for	TCGTCCATACAGCACTCGTC	rsp-7 PTC(+) rev 1	GGTAGTGGCCAGTAGAACCAAAG
rsp-7(+)	rsp-7 e3 for	TCGTCCATACAGCACTCGTC	rsp-7 PTC(-) rev 1	GTGTCGGATCGACTGTAATCAACAAAG
rsp-7(total)	rsp-7 e2/e3 for	CCTACCATATCCGAATCCTGTG	rsp-7 e3 rev	GTGAGGTGGGAGCTGTCGC
eft-3	eft-3 F	TATCAACATCGTCGTCATCG	eft-3 R	CCTCCTTCTCGAACTTCTCG
rsp-2(endogenous)	rsp-2 PTC(-) for 2	CAGTTGGCAGGACATCAAGGAC	rsp-2 3' UTR rev 2	GAGAGACGCACAACATTAGATT
rsp-4(endogenous)	rsp-4 PTC(-) endo for	GTGGTGGTGGACGTAGAAGATCC	rsp-4 3' UTR rev 2	TATTACTTGTCTTGTTGGAGGC
rsp-6(endogenous)	rsp-6 PTC(-) endo for	GTGGCGGAAGAGACAGATCG	rsp-6 3' UTR rev 1	TTGATGTTGTGGAGAAGATACTG

Table M.4. Primers to generate *gfp::rsp* transgenes and express recombinant RSP proteins in *E. coli* (for antibody production/purification).

Product	Forward Primer	Forward Primer Sequence	Reverse Primer	Reverse Primer Sequence
rsp-4 gDNA and rsp-4 fusion product	rsp-4 For Cut (NheI)	TATAGCTAGCATTTCAGAAATGAGCCGT	rsp-4 Rev Cut (NcoI)	TATACCATGGTAACTGCTTTGACACTTGA
rsp-4 fusion 1	rsp-4 For Cut (NheI)	TATAGCTAGCATTTCAGAAATGAGCCGT	rsp-4 e2 (fus-1)	GAGCGGGATCTTCTACGTCCACCACC
rsp-4 fusion 2	rsp-4 e4 (fus-1)	GACGTAGAAGATCCCGCTCACCGAGAAG	rsp-4 Rev Cut (NcoI)	TATACCATGGTAACTGCTTTGACACTTGA
rsp-4 Final Vector	rsp-4 For (AgeI)	TATAACCGGTATTTCAGAAATGAGCCGT	rsp-4 Rev Cut (NcoI) STOP	TAAACCATGGTTAGCGGGAGTTGGAAC
rsp-2 gDNA	rsp-2 for cut	TAAAGAATTCAGACCTCTAAACATGGTTC	rsp-2 rev cut (NcoI) STOP	TAAACCATGGTTATGGAGATGCAGATCCG
rsp-2 fusion 1	rsp-2 for cut	TAAAGAATTCAGACCTCTAAACATGGTTC	r2e3 (fus-1)	GGATGTGGTCCTTGATGTCCTCCTGCCAACT
rsp-2 fusion 2	r2e2 (fus-2)	CTACTCGTTACAGTTGGCAGGAGGACATCAAGGACCAC	rsp-2 rev cut (NcoI) STOP	TAAACCATGGTTATGGAGATGCAGATCCG
rsp-6 gDNA and fusion product	rsp-6 for cut(new)	TAAAGACGTCATTTGATTTTTCAGTATG	rsp-6 rev cut (NcoI) STOP	TAAACCATGGTTAGTGCGGAGAAGCAGAAC
rsp-6 fusion 1	rsp-6 for cut(new)	TAAAGACGTCATTTGATTTTTCAGTATG	rsp-6 fus-2	GTATGGCGATCTGTCTCTTCCGCCACC
rsp-6 fusion 2	rsp-6 fus-1	GAAGAGACAGATCGCCATACCGAGGAG	rsp-6 rev cut (NcoI) STOP	TAAACCATGGTTAGTGCGGAGAAGCAGAAC
GFP for rsp-2 and rsp-6	GFP for cut(AgeI)	TAAAACCGGTCCGGTAGAAAAAATGAGTA	GFP rev cut (New)	TAAAGACGTCTACGAATGCAATTTGTATAGTTC
GFP for rsp-4	GFP for cut(AgeI)	TAAAACCGGTCCGGTAGAAAAAATGAGTA	GFP Rev Cut (NheI) - FS	TAAAGCTAGCATACGAATGCAATTTGTATAGTTC
GST::RSP-2	rsp-2 for (ATG) pGEX-4T-1 EcoRI	CCGGGAGCTGCATGTGTCAGAGG	rsp-2 rev (e2) pGEX-4T-1 XhoI	TATACTCGAGCCTGCCAACTGTAACG
GST::RSP-6	rsp-6 for (ATG) pGEX-4T-1 EcoRI	TATAGAATTCATGGACGCCAAGGTG	rsp-6 rev (e2) pGEX-4T-1 XhoI	TATACTCGAGGTCAAGAGCGCGG
GST::RSP-4	rsp-4 for (ATG) pGEX-4T-1 EcoRI	TATAGAATTCATGAGCCGTGGAGGAG	rsp-4 rev (e2) pGEX-4T-1 XhoI	TATACTCGAGCTTCTACGTCCACC
His::RSP-2	rsp-2 for(ATG)pET15-b NdeI	TATACATATGATGGTTCGTGTCTACATC	rsp-2 rev(TAA) pET-15b XhoI	TATACTCGAGTTATGGAGATGCAGATC
His::RSP-6	rsp-6 for(ATG)pET15-b NdeI	TATACATATGATGGACGCCAAGGTG	rsp-6 rev (TAA) pET-15b XhoI	TATACTCGAGTTAGTGCGGAGAAGC
MBP::RSP-2	rsp-2 for(ATG) pGEX-4T-1 EcoRI	CCGGGAGCTGCATGTGTCAGAGG	pMal-c2x rsp-2 short rev	TATAAAGCTTCCTACCAACTGTAACG
MBP::RSP-6	rsp-6 for(ATG) pGEX-4T-1 EcoRI	TATAGAATTCATGGACGCCAAGGTG	pMal-c2x rsp-6 long rev	TATAAAGCTTTTAGTGCGGAGAAGC

References

1. Black DL: **Mechanisms of alternative pre-messenger RNA splicing**. *Annu Rev Biochem* 2003, **72**:291–336.

2. Nilsen TW, Graveley BR: **Expansion of the eukaryotic proteome by alternative splicing.** *Nature* 2010, **463**:457–463.

3. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ: **Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing.** *Nature Genetics, Published online: 11 February 2007; | doi:10.1038/ng1983* 2008, **40**:1413–1415.

4. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB: **Alternative isoform regulation in human tissue transcriptomes.** *Nature* 2008, **456**:470–476.

5. Kalsotra A, Cooper TA: **Functional consequences of developmentally regulated alternative splicing**. *Nat Rev Genet* 2011, **12**:715–729.

6. Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al.: **The developmental transcriptome of Drosophila melanogaster**. *Nature* 2011, **471**:473–479.

7. Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, Wong W-K, Mockler TC: **Genome-wide mapping of alternative splicing in Arabidopsis thaliana.** *Genome Res* 2010, **20**:45–58.

8. Ramani AK, Calarco JA, Pan Q, Mavandadi S, Wang Y, Nelson AC, Lee LJ, Morris Q, Blencowe BJ, Zhen M, et al.: **Genome-wide analysis of alternative splicing in Caenorhabditis elegans.** *Genome Res* 2011, **21**:342–348.

9. Mayeda A, Krainer AR: **Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2.** *Cell* 1992, **68**:365–375.

10. Howe KJ: **RNA polymerase II conducts a symphony of pre-mRNA processing activities.** *Biochim Biophys Acta* 2002, **1577**:308–324.

11. la Mata de M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR: **A slow RNA polymerase II affects alternative splicing in vivo.** *Mol Cell* 2003, **12**:525–532.

12. Ringrose L: **How do RNA sequence, DNA sequence, and chromatin properties regulate splicing?** *F1000 Biol Rep* 2010, **2**:74.

13. Carrillo Oesterreich F, Bieberstein N, Neugebauer KM: **Pause locally, splice globally.** *Trends Cell Biol.* 2011, **21**:328–335.

14. Blencowe BJ, Bowman JA, McCracken S, Rosonina E: **SR-related proteins and the processing of messenger RNA precursors**. *Biochem Cell Biol* 1999, **77**:277–291.

15. Graveley BR, Hertel KJ, Maniatis T: **SR proteins are "locators" of the RNA splicing machinery**. *Curr Biol* 1999, **9**:R6–7.

16. Tacke R, Manley JL: **Determinants of SR protein specificity.** *Curr Opin Cell Biol* 1999, **11**:358–362.

17. Graveley BR: **Sorting out the complexity of SR protein functions**. *RNA* 2000, **6**:1197–1211.

18. Sanford JR, Longman D, Cáceres JF: **Multiple roles of the SR protein family in splicing regulation**. *Prog Mol Subcell Biol* 2003, **31**:33–58.

19. Bourgeois CF, Lejeune F, Stévenin J: **Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA**. *Prog Nucleic Acid Res Mol Biol* 2004, **78**:37–88.

20. Sanford JR, Ellis J, Cáceres JF: **Multiple roles of arginine/serine-rich splicing factors in RNA processing**. *Biochem Soc Trans* 2005, **33**:443–446.

21. Shepard PJ, Hertel KJ: **The SR protein family**. *Genome Biol* 2009, **10**:242.

22. Busch A, Hertel KJ: **Evolution of SR protein and hnRNP splicing regulatory factors**. *Wiley Interdiscip Rev RNA* 2011, doi:10.1002/wrna.100.

23. Chen M, Manley JL: **Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches.** *Nat Rev Mol Cell Biol* 2009, **10**:741–754.

24. Long JC, Cáceres JF: **The SR protein family of splicing factors: master regulators of gene expression.** *Biochem. J.* 2009, **417**:15–27.

25. Zahler AM, Lane WS, Stolk JA, Roth MB: **SR proteins: a conserved family of pre-mRNA splicing factors**. *Genes Dev* 1992, **6**:837–847.

26. Zahler AM, Neugebauer KM, Lane WS, Roth MB: **Distinct functions of SR proteins in alternative pre-mRNA splicing**. *Science* 1993, **260**:219–222.

27. Zahler AM, Neugebauer KM, Stolk JA, Roth MB: **Human SR proteins and isolation of a cDNA encoding SRp75**. *Mol Cell Biol* 1993, **13**:4023–4028.

28. Krainer AR, Conway GC, Kozak D: **Purification and characterization of pre-**

mRNA splicing factor SF2 from HeLa cells. *Genes Dev* 1990, **4**:1158–1171.

29. Krainer AR, Mayeda A, Kozak D, Binns G: **Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators.** *Cell* 1991, **66**:383–394.

30. Mayeda A, Zahler AM, Krainer AR, Roth MB: **Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing.** *Proc Natl Acad Sci USA* 1992, **89**:1301–1304.

31. Gopinath SCB: **Methods developed for SELEX**. *Analytical and bioanalytical chemistry* 2007, **387**:171–182.

32. Kim S, Shi H, Lee D-K, Lis JT: **Specific SR protein-dependent splicing substrates identified through genomic SELEX**. *Nucleic Acids Res* 2003, **31**:1955–1961.

33. Wang J, Smith PJ, Krainer AR, Zhang MQ: **Distribution of SR protein exonic splicing enhancer motifs in human protein-coding genes.** *Nucleic Acids Res* 2005, **33**:5053–5062.

34. Wu JY, Maniatis T: **Specific interactions between proteins implicated in splice site selection and regulated alternative splicing.** *Cell* 1993, **75**:1061–1070.

35. Kohtz JD, Jamison SF, Will CL, Zuo P, Lührmann R, Garcia-Blanco MA, Manley JL: **Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors.** *Nature* 1994, **368**:119–124.

36. Zahler AM, Roth MB: **Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites.** *Proc Natl Acad Sci USA* 1995, **92**:2642–2646.

37. Manley JL, Krainer AR: **A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins).** *Genes Dev* 2010, **24**:1073–1074.

38. Barta A, Kalyna M, Reddy ASN: **Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors (SR proteins) in plants.** *Plant Cell* 2010, **22**:2926–2929.

39. Palusa SG, Ali GS, Reddy ASN: **Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses.** *Plant J.* 2007, **49**:1091–1107.

40. Kress TL, Krogan NJ, Guthrie C: **A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast.** *Mol Cell* 2008, **32**:727–734.

41. Longman D, Johnstone IL, Cáceres JF: **Functional characterization of SR and SR-**

related genes in Caenorhabditis elegans. *EMBO J* 2000, **19**:1625–1637.

42. Kawano T, Fujita M, Sakamoto H: **Unique and redundant functions of SR proteins, a conserved family of splicing factors, in Caenorhabditis elegans development.** *Mech Dev* 2000, **95**:67–76.

43. Wang J, Takagaki Y, Manley JL: **Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability.** *Genes Dev* 1996, **10**:2588–2599.

44. Richardson DN, Rogers MF, Labadorf A, Ben-Hur A, Guo H, Paterson AH, Reddy ASN: **Comparative analysis of serine/arginine-rich proteins across 27 eukaryotes: insights into sub-family classification and extent of alternative splicing.** *PLoS ONE* 2011, **6**:e24542.

45. Morrison M, Harris KS, Roth MB: **smg mutants affect the expression of alternatively spliced SR protein mRNAs in Caenorhabditis elegans.** *Proc Natl Acad Sci USA* 1997, **94**:9782–9785.

46. Duque P: **A role for SR proteins in plant stress responses.** *Plant Signal Behav* 2011, **6**:49–54.

47. Lareau LF, Brooks AN, Soergel DAW, Meng Q, Brenner SE: **The coupling of alternative splicing and nonsense-mediated mRNA decay**. *Adv Exp Med Biol* 2007, **623**:190–211.

48. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE: **Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements.** *Nature* 2007, **446**:926–929.

49. Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares M: **Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay**. *Genes Dev* 2007, **21**:708–718.

50. Anko M-L, Muller-McNicoll M, Brandl H, Curk T, Gorup C, Henry I, Ule J, Neugebauer KM: **The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes.** *Genome Biol* 2012, **13**:R17.

51. Tanabe N, Yoshimura K, Kimura A, Yabuta Y, Shigeoka S: **Differential Expression of Alternatively Spliced mRNAs of Arabidopsis SR Protein Homologs, atSR30 and atSR45a, in Response to Environmental Stress**. *Plant and Cell Physiology* 2007, **48**:1036–1049.

52. Ip JY, Schmidt D, Pan Q, Ramani AK, Fraser AG, Odom DT, Blencowe BJ: **Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation.** *Genome Res* 2011, **21**:390–401.

53. Baugh LR, Demodena J, Sternberg PW: **RNA Pol II accumulates at promoters of growth genes during developmental arrest.** *Science* 2009, **324**:92–94.

54. Sanford JR, Bruzik JP: **Developmental regulation of SR protein phosphorylation and activity**. *Genes Dev* 1999, **13**:1513–1518.

55. Stamm S: **Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome.** *Hum. Mol. Genet.* 2002, **11**:2409–2416.

56. Guil S, Cáceres JF: **Stressful splicing**. *Mol Cell* 2007, **28**:180–181.

57. Shin C, Manley JL: **Cell signalling and the control of pre-mRNA splicing.** *Nat Rev Mol Cell Biol* 2004, **5**:727–738.

58. Dreyfuss G, Matunis MJ, Piñol-Roma S, Burd CG: **hnRNP proteins and the biogenesis of mRNA.** *Annu Rev Biochem* 1993, **62**:289–321.

59. Görlach M, Burd CG, Portman DS, Dreyfuss G: **The hnRNP proteins.** *Mol. Biol. Rep.* 1993, **18**:73–78.

60. Han SP, Tang YH, Smith R: **Functional diversity of the hnRNPs: past, present and perspectives.** *Biochem. J.* 2010, **430**:379–392.

61. Dreyfuss G, Choi YD, Adam SA: **Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies.** *Mol Cell Biol* 1984, **4**:1104–1114.

62. Dreyfuss G: **Structure and function of nuclear and cytoplasmic ribonucleoprotein particles.** *Annu. Rev. Cell Biol.* 1986, **2**:459–498.

63. Choi YD, Dreyfuss G: **Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly.** *Proc Natl Acad Sci USA* 1984, **81**:7471–7475.

64. Piñol-Roma S, Choi YD, Matunis MJ, Dreyfuss G: **Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins.** *Genes Dev* 1988, **2**:215–227.

65. Choi YD, Grabowski PJ, Sharp PA, Dreyfuss G: **Heterogeneous nuclear ribonucleoproteins: role in RNA splicing.** *Science* 1986, **231**:1534–1539.

66. Sierakowska H, Szer W, Furdon PJ, Kole R: **Antibodies to hnRNP core proteins inhibit in vitro splicing of human beta-globin pre-mRNA.** *Nucleic Acids Res* 1986, **14**:5241–5254.

67. Mayeda A, Helfman DM, Krainer AR: **Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF.** *Mol Cell Biol* 1993, **13**:2993–3001.

68. Blanchette M, Green RE, MacArthur S, Brooks AN, Brenner SE, Eisen MB, Rio DC: **Genome-wide analysis of alternative pre-mRNA splicing and RNA-binding specificities of the Drosophila hnRNP A/B family members.** *Mol Cell* 2009, **33**:438–449.

69. Kornblihtt AR: **Coupling transcription and alternative splicing.** *Adv Exp Med Biol* 2007, **623**:175–189.

70. Allemand E, Batsché E, Muchardt C: **Splicing, transcription, and chromatin: a ménage à trois.** *Curr Opin Genet Dev* 2008, **18**:145–151.

71. Kornblihtt AR: **Chromatin, transcript elongation and alternative splicing**. *Nat Struct Mol Biol* 2006, **13**:5–7.

72. Moore MJ, Proudfoot NJ: **Pre-mRNA processing reaches back to transcription and ahead to translation.** *Cell* 2009, **136**:688–700.

73. Das R, Yu J, Zhang Z, Gygi MP, Krainer AR, Gygi SP, Reed R: **SR proteins function in coupling RNAP II transcription to pre-mRNA splicing.** *Mol Cell* 2007, **26**:867–881.

74. la Mata de M, Kornblihtt AR: **RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20.** *Nat Struct Mol Biol* 2006, **13**:973–980.

75. Schwartz S, Ast G: **Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing.** *EMBO J* 2010, **29**:1629–1636.

76. Shukla S, Oberdoerffer S: **Co-transcriptional regulation of alternative pre-mRNA splicing.** *Biochim Biophys Acta* 2012, doi:10.1016/j.bbagrm.2012.01.014.

77. Alló M, Schor IE, Muñoz MJ, La Mata De M, Agirre E, Valcárcel J, Eyras E, Kornblihtt AR: **Chromatin and alternative splicing.** *Cold Spring Harb. Symp. Quant. Biol.* 2010, **75**:103–111.

78. Kouzarides T: **Chromatin modifications and their function.** *Cell* 2007, **128**:693–705.

79. Schones DE, Cui K, Cuddapah S, Roh T-Y, Barski A, Wang Z, Wei G, Zhao K: **Dynamic regulation of nucleosome positioning in the human genome.** *Cell* 2008, **132**:887–898.

80. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek

JA, Costa G, McKernan K, et al.: **A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning.** *Genome Res* 2008, **18**:1051–1063.

81. Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcárcel J, Guigó R: **Nucleosome positioning as a determinant of exon recognition.** *Nat Struct Mol Biol* 2009, **16**:996–1001.

82. Spies N, Nielsen CB, Padgett RA, Burge CB: **Biased chromatin signatures around polyadenylation sites and exons.** *Mol Cell* 2009, **36**:245–254.

83. Schwartz S, Meshorer E, Ast G: **Chromatin organization marks exon-intron structure.** *Nat Struct Mol Biol* 2009, **16**:990–995.

84. Chen W, Luo L, Zhang L: **The organization of nucleosomes around splice sites.** *Nucleic Acids Res* 2010, **38**:2788–2798.

85. Schor IE, Alló M, Kornblihtt AR: **Intragenic chromatin modifications: A new layer in alternative splicing regulation.** *Epigenetics* 2010, **5**.

86. Luco RF, Alló M, Schor IE, Kornblihtt AR, Misteli T: **Epigenetics in alternative pre-mRNA splicing.** *Cell* 2011, **144**:16–26.

87. Irimia M, Blencowe BJ: **Alternative splicing: decoding an expansive regulatory layer.** *Curr Opin Cell Biol* 2012, doi:10.1016/j.ceb.2012.03.005.

88. Greer EL, Shi Y: **Histone methylation: a dynamic mark in health, disease and inheritance.** *Nat Rev Genet* 2012, doi:10.1038/nrg3173.

89. Bell O, Wirbelauer C, Hild M, Scharf AND, Schwaiger M, MacAlpine DM, Zilbermann F, van Leeuwen F, Bell SP, Imhof A, et al.: **Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila.** *EMBO J* 2007, **26**:4974–4984.

90. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J: **Nucleosomes are well positioned in exons and carry characteristic histone modifications.** *Genome Res* 2009, **19**:1732–1741.

91. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J: **Differential chromatin marking of introns and expressed exons by H3K36me3.** *Nature Genetics, Published online: 11 February 2007; | doi:10.1038/ng1983* 2009, **41**:376–381.

92. Sims RJ, Reinberg D: **Processing the H3K36me3 signature.** *Nature Genetics, Published online: 11 February 2007; | doi:10.1038/ng1983* 2009, **41**:270–271.

93. Hon G, Wang W, Ren B: **Discovery and annotation of functional chromatin signatures in the human genome.** *PLoS Comput Biol* 2009, **5**:e1000566.

94. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T: **Regulation of alternative splicing by histone modifications.** *Science* 2010, **327**:996–1000.

95. Wagner EJ, Carpenter PB: **Understanding the language of Lys36 methylation at histone H3.** *Nat Rev Mol Cell Biol* 2012, **13**:115–126.

96. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S: **CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing**. *Nature* 2011, **479**:74–79.

97. Kornblihtt AR: **CTCF: from insulators to alternative splicing regulation.** *Cell Res.* 2012, **22**:450–452.

98. Alló M, Buggiano V, Fededa JP, Petrillo E, Schor I, la Mata de M, Agirre E, Plass M, Eyras E, Elela SA, et al.: **Control of alternative splicing through siRNA-mediated transcriptional gene silencing.** *Nat Struct Mol Biol* 2009, **16**:717–724.

99. Ule J, Ule A, Spencer J, Williams A, Hu J-S, Cline M, Wang H, Clark T, Fraser C, Ruggiu M, et al.: **Nova regulates brain-specific splicing to shape the synapse.** *Nature Genetics, Published online: 11 February 2007; | doi:10.1038/ng1983* 2005, **37**:844–852.

100. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, et al.: **HITS-CLIP yields genome-wide insights into brain alternative RNA processing.** *Nature* 2008, **456**:464–469.

101. Zhang C, Frias MA, Mele A, Ruggiu M, Eom T, Marney CB, Wang H, Licatalosi DD, Fak JJ, Darnell RB: **Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls.** *Science* 2010, **329**:439–443.

102. Boutz PL, Stoilov P, Li Q, Lin C-H, Chawla G, Ostrow K, Shiue L, Ares M, Black DL: **A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons.** *Genes Dev* 2007, **21**:1636–1652.

103. Makeyev EV, Zhang J, Carrasco MA, Maniatis T: **The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing.** *Mol Cell* 2007, **27**:435–448.

104. Spellman R, Llorian M, Smith CWJ: **Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1.** *Mol Cell* 2007, **27**:420–434.

105. Llorian M, Schwartz S, Clark TA, Hollander D, Tan L-Y, Spellman R, Gordon A, Schweitzer AC, la Grange de P, Ast G, et al.: **Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB.** *Nat Struct Mol Biol* 2010, **17**:1114–1123.

106. Isken O, Maquat LE: **The multiple lives of NMD factors: balancing roles in gene and genome regulation.** *Nat Rev Genet* 2008, **9**:699–712.

107. Isken O, Maquat LE: **Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function.** *Genes Dev* 2007, **21**:1833–1856.

108. Behm-Ansmant I, Kashima I, Rehwinkel J, Saulière J, Wittkopp N, Izaurralde E: **mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons.** *FEBS Lett* 2007, **581**:2845–2853.

109. Chang Y-F, Imam JS, Wilkinson MF: **The nonsense-mediated decay RNA surveillance pathway.** *Annu Rev Biochem* 2007, **76**:51–74.

110. Behm-Ansmant I, Izaurralde E: **Quality control of gene expression: a stepwise assembly pathway for the surveillance complex that triggers nonsense-mediated mRNA decay.** *Genes Dev* 2006, **20**:391–398.

111. Weischenfeldt J, Lykke-Andersen J, Porse B: **Messenger RNA surveillance: neutralizing natural nonsense.** *Curr Biol* 2005, **15**:R559–62.

112. Conti E, Izaurralde E: **Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species.** *Curr Opin Cell Biol* 2005, **17**:316–325.

113. Maquat LE: **Nonsense-mediated mRNA decay in mammals.** *J Cell Sci* 2005, **118**:1773–1776.

114. Maquat LE: **Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics**. *Nat Rev Mol Cell Biol* 2004, **5**:89–99.

115. Culbertson MR, Leeds PF: **Looking at mRNA decay pathways through the window of molecular evolution.** *Curr Opin Genet Dev* 2003, **13**:207–214.

116. Nicholson P, Yepiskoposyan H, Metze S, Zamudio Orozco R, Kleinschmidt N, Mühlemann O: **Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors.** *Cell Mol Life Sci* 2010, **67**:677–700.

117. Hwang J, Maquat LE: **Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question.** *Curr Opin Genet Dev* 2011, **21**:422–430.

118. Schoenberg DR, Maquat LE: **Regulation of cytoplasmic mRNA decay.** *Nat Rev Genet* 2012, **13**:246–259.

119. Culbertson MR, Underbrink KM, Fink GR: **Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors.** *Genetics* 1980, **95**:833–853.

120. Leeds P, Peltz SW, Jacobson A, Culbertson MR: **The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon.** *Genes Dev* 1991, **5**:2303–2314.

121. Cui Y, Hagan KW, Zhang S, Peltz SW: **Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon.** *Genes Dev* 1995, **9**:423–436.

122. Hodgkin J, Papp A, Pulak R, Ambros V, Anderson P: **A new kind of informational suppression in the nematode Caenorhabditis elegans**. *Genetics* 1989, **123**:301–313.

123. Pulak R, Anderson P: **mRNA surveillance by the Caenorhabditis elegans smg genes.** *Genes Dev* 1993, **7**:1885–1897.

124. Grimson A, O'Connor S, Newman CL, Anderson P: **SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans.** *Mol Cell Biol* 2004, **24**:7483–7490.

125. Page MF, Carr B, Anders KR, Grimson A, Anderson P: **SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast.** *Mol Cell Biol* 1999, **19**:5943–5951.

126. Anders KR, Grimson A, Anderson P: **SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A.** *EMBO J* 2003, **22**:641–650.

127. Aronoff R, Baran R, Hodgkin J: **Molecular identification of smg-4, required for mRNA surveillance in C. elegans.** *Gene* 2001, **268**:153–164.

128. Cali BM, Kuchma SL, Latham J, Anderson P: **smg-7 is required for mRNA surveillance in Caenorhabditis elegans.** *Genetics* 1999, **151**:605–616.

129. Gatfield D, Unterholzner L, Ciccarelli FD, Bork P, Izaurralde E: **Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways.** *EMBO J* 2003, **22**:3960–3970.

130. Longman D, Plasterk RHA, Johnstone IL, Cáceres JF: **Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway.** *Genes Dev*

2007, **21**:1075–1085.

131. Yamashita A, Izumi N, Kashima I, Ohnishi T, Saari B, Katsuhata Y, Muramatsu R, Morita T, Iwamatsu A, Hachiya T, et al.: **SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay.** *Genes Dev* 2009, **23**:1091–1105.

132. Yamashita A, Ohnishi T, Kashima I, Taya Y, Ohno S: **Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay.** *Genes Dev* 2001, **15**:2215–2228.

133. Johns L, Grimson A, Kuchma SL, Newman CL, Anderson P: **Caenorhabditis elegans SMG-2 selectively marks mRNAs containing premature translation termination codons.** *Mol Cell Biol* 2007, **27**:5630–5638.

134. Johansson MJO, He F, Spatrick P, Li C, Jacobson A: **Association of yeast Upf1p with direct substrates of the NMD pathway.** *Proc Natl Acad Sci USA* 2007, **104**:20872–20877.

135. Hogg JR, Goff SP: **Upf1 senses 3'UTR length to potentiate mRNA decay.** *Cell* 2010, **143**:379–389.

136. Kashima I, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, Ohno M, Dreyfuss G, Ohno S: **Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay.** *Genes Dev* 2006, **20**:355–367.

137. Czaplinski K, Ruiz-Echevarria MJ, Paushkin SV, Han X, Weng Y, Perlick HA, Dietz HC, Ter-Avanesyan MD, Peltz SW: **The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs.** *Genes Dev* 1998, **12**:1665–1677.

138. Mühlemann O, Lykke-Andersen J: **How and where are nonsense mRNAs degraded in mammalian cells?** *RNA Biol* 2010, **7**:28–32.

139. Huntzinger E, Kashima I, Fauser M, Saulière J, Izaurralde E: **SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan.** *RNA* 2008, **14**:2609–2617.

140. Eberle AB, Lykke-Andersen S, Mühlemann O, Jensen TH: **SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells.** *Nat Struct Mol Biol* 2009, **16**:49–55.

141. Amrani N, Sachs MS, Jacobson A: **Early nonsense: mRNA decay solves a translational problem.** *Nat Rev Mol Cell Biol* 2006, **7**:415–425.

142. Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, Jacobson A: **A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay.** *Nature* 2004, **432**:112–118.

143. Bühler M, Steiner S, Mohn F, Paillusson A, Mühlemann O: **EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length.** *Nat Struct Mol Biol* 2006, **13**:462–464.

144. Behm-Ansmant I, Gatfield D, Rehwinkel J, Hilgers V, Izaurralde E: **A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay.** *EMBO J* 2007, **26**:1591–1601.

145. Singh G, Rebbapragada I, Lykke-Andersen J: **A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay.** *PLoS Biol* 2008, **6**:e111.

146. Cali BM, Anderson P: **mRNA surveillance mitigates genetic dominance in Caenorhabditis elegans.** *Mol. Gen. Genet.* 1998, **260**:176–184.

147. Hansen KD, Lareau LF, Blanchette M, Green RE, Meng Q, Rehwinkel J, Gallusser FL, Izaurralde E, Rio DC, Dudoit S, et al.: **Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in Drosophila.** *PLoS Genet* 2009, **5**:e1000525.

148. Kurihara Y, Matsui A, Hanada K, Kawashima M, Ishida J, Morosawa T, Tanaka M, Kaminuma E, Mochizuki Y, Matsushima A, et al.: **Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis.** *Proc Natl Acad Sci USA* 2009, **106**:2453–2458.

149. Ramani AK, Nelson AC, Kapranov P, Bell I, Gingeras TR, Fraser AG: **High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans.** *Genome Biol* 2009, **10**:R101.

150. Sayani S, Janis M, Lee CY, Toesca I, Chanfreau GF: **Widespread impact of nonsense-mediated mRNA decay on the yeast intronome**. *Mol Cell* 2008, **31**:360–370.

151. Saltzman AL, Kim YK, Pan Q, Fagnani MM, Maquat LE, Blencowe BJ: **Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay.** *Mol Cell Biol* 2008, **28**:4320–4330.

152. Guan Q, Zheng W, Tang S, Liu X, Zinkel RA, Tsui K-W, Yandell BS, Culbertson MR: **Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast.** *PLoS Genet* 2006, **2**:e203.

153. Rehwinkel J, Letunic I, Raes J, Bork P, Izaurralde E: **Nonsense-mediated mRNA**

decay factors act in concert to regulate common mRNA targets. *RNA* 2005, **11**:1530–1544.

154. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC: **Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise.** *Nature Genetics, Published online: 11 February 2007; | doi:10.1038/ng1983* 2004, **36**:1073–1078.

155. He F, Li X, Spatrick P, Casillo R, Dong S, Jacobson A: **Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5“ to 3” mRNA decay pathways in yeast.** *Mol Cell* 2003, **12**:1439–1452.

156. Mitrovich QM, Anderson P: **mRNA surveillance of expressed pseudogenes in C. elegans.** *Curr Biol* 2005, **15**:963–967.

157. Lewis BP, Green RE, Brenner SE: **Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans.** *Proc Natl Acad Sci USA* 2003, **100**:189–192.

158. Green RE, Lewis BP, Hillman RT, Blanchette M, Lareau LF, Garnett AT, Rio DC, Brenner SE: **Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes.** *Bioinformatics* 2003, **19 Suppl 1**:i118–21.

159. Mitrovich QM, Anderson P: **Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans**. *Genes Dev* 2000, **14**:2173–2184.

160. Cuccurese M, Russo G, Russo A, Pietropaolo C: **Alternative splicing and nonsense-mediated mRNA decay regulate mammalian ribosomal gene expression.** *Nucleic Acids Res* 2005, **33**:5965–5977.

161. Sureau A, Gattoni R, Dooghe Y, Stévenin J, Soret J: **SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs.** *EMBO J* 2001, **20**:1785–1796.

162. Rossbach O, Hung L-H, Schreiner S, Grishina I, Heiner M, Hui J, Bindereif A: **Auto- and cross-regulation of the hnRNP L proteins by alternative splicing.** *Mol Cell Biol* 2009, **29**:1442–1451.

163. Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CWJ: **Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay.** *Mol Cell* 2004, **13**:91–100.

164. Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, Maquat LE, Frey BJ, Blencowe BJ: **Quantitative microarray profiling provides evidence against widespread**

coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes Dev* 2006, **20**:153–158.

165. McGlincy NJ, Smith CWJ: **Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense?** *Trends Biochem Sci* 2008, **33**:385–393.

166. Soller M: **Pre-messenger RNA processing and its regulation: a genomic perspective.** *Cell Mol Life Sci* 2006, **63**:796–819.

167. Mudge JM, Frankish A, Fernandez-Banet J, Alioto T, Derrien T, Howald C, Reymond A, Guigó R, Hubbard T, Harrow J: **The origins, evolution, and functional potential of alternative splicing in vertebrates.** *Mol. Biol. Evol.* 2011, **28**:2949–2959.

168. Salz HK: **Sex determination in insects: a binary decision based on alternative splicing**. *Curr Opin Genet Dev* 2011, **21**:395–400.

169. Sánchez L: **Sex-determining mechanisms in insects**. *Int J Dev Biol* 2008, **52**:837–856.

170. Zahler AM: **Alternative splicing in C. elegans.** *WormBook : the online review of C elegans biology* 2005, doi:10.1895/wormbook.1.31.1.

171. Barberan-Soler S, Zahler AM: **Alternative splicing regulation during C. elegans development: splicing factors as regulated targets.** *PLoS Genet* 2008, **4**:e1000001.

172. Palusa SG, Reddy ASN: **Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay.** *New Phytol.* 2010, **185**:83–89.

173. Kim YK, Furic L, Desgroseillers L, Maquat LE: **Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay.** *Cell* 2005, **120**:195–208.

174. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I: **VISTA: computational tools for comparative genomics.** *Nucleic Acids Res* 2004, **32**:W273–9.

175. Jiang M, Ryu J, Kiraly M, Duke K, Reinke V, Kim SK: **Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans.** *Proc Natl Acad Sci USA* 2001, **98**:218–223.

176. Blackwell TK, Walker AK: **Transcription mechanisms.** *WormBook : the online review of C elegans biology* 2006, doi:10.1895/wormbook.1.121.1.

177. Edgley ML, Baillie DL, Riddle DL, Rose AM: **Genetic balancers.** *WormBook : the*

online review of C elegans biology 2006, doi:10.1895/wormbook.1.89.1.

178. Merritt C, Seydoux G: **Transgenic solutions for the germline.** *WormBook : the online review of C elegans biology* 2010, doi:10.1895/wormbook.1.148.1.

179. Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S: **Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription.** *Nature* 2010, **465**:1097–1101.

180. Ryder SP, Frater LA, Abramovitz DL, Goodwin EB, Williamson JR: **RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1.** *Nat Struct Mol Biol* 2004, **11**:20–28.

181. Sambrook J, Russell DW: *Molecular Cloning*. CSHL Press; 2001.