

---

# Alternative splicing in *C. elegans*\*

Alan M. Zahler<sup>§</sup>, Department of Molecular, Cellular and Developmental Biology, Center for Molecular Biology of RNA, Sinsheimer Laboratories, University of California, Santa Cruz, CA 95064 USA

## Table of Contents

|   |   |
|---|---|
| 1. The Importance of alternative splicing .....                             | 1 |
| 2. Splicing of precursors to messenger RNA .....                            | 3 |
| 3. Models for alternative splicing regulation .....                         | 3 |
| 4. Constitutive splicing factors in <i>C. elegans</i> .....                 | 4 |
| 5. Identification of alternatively spliced genes in <i>C. elegans</i> ..... | 5 |
| 6. Some examples of alternatively spliced <i>C. elegans</i> genes .....     | 5 |
| 7. Regulation of alternative splicing in <i>C. elegans</i> .....            | 6 |
| 8. Conclusions .....  | 8 |
| 9. Acknowledgements .....   | 8 |
| 10. References .....  | 8 |

## Abstract

Alternative splicing is a common mechanism for the generation of multiple isoforms of proteins. It can function to expand the proteome of an organism and can serve as a way to turn off gene expression post-transcriptionally. This review focuses on splicing and its regulation in *C. elegans*. The fully-sequenced *C. elegans* genome combined with its elegant genetics offers unique advantages for exploring alternative splicing regulation in metazoans. The topics covered in this review include constitutive splicing factors, identification of alternatively spliced genes, examples of alternative splicing in *C. elegans*, and alternative splicing regulation. Key genes whose regulated alternative splicing are reviewed include *let-2*, *unc-32*, *unc-52*, *egl-15* and *xol-1*. Factors involved in alternative splicing that are discussed include *mec-8*, *smu-1*, *smu-2*, *fox-1*, *exc-7* and *unc-75*.

## 1. The Importance of alternative splicing

Advances in animal genome sequencing have led to many questions concerning the very nature and number of genes and how they help to promote the diversity of cellular and organismal functions. One of the interesting disappointments of the analysis of the human genome sequence is that we may possess less than 25,000 genes (International Human Genome Sequencing Consortium, 2004), only one third more than the number of predicted

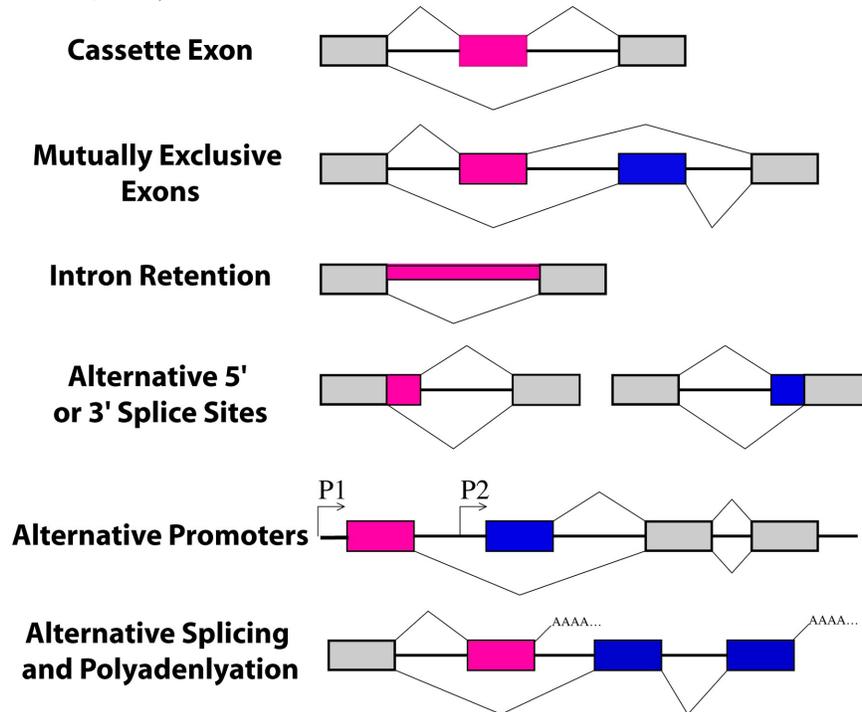
---

\*Edited by Thomas Blumenthal. Last revised January 28, 2005. Published September 26, 2005. This chapter should be cited as: Zahler, A.M. Alternative splicing in *C. elegans* (September 26, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.31.1, <http://www.wormbook.org>.

**Copyright:** © 2005 Alan M. Zahler. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<sup>§</sup>To whom correspondence should be addressed. E-mail: [zahler@biology.ucsc.edu](mailto:zahler@biology.ucsc.edu)

genes in *C. elegans*. How can we explain all the complexity and wonder of human biology with such a limited gene set? One way to explain this paradox is to point out that the number of possible proteins from the genome can far exceed the possible number of genes if a large percentage of the genes have the ability to encode multiple proteins. This expansion of the proteome can be accomplished through alternative precursor messenger RNA (pre-mRNA) splicing, which can allow one gene to encode multiple proteins. Comparison of cDNAs with genomic sequence has provided evidence for extensive alternative splicing of human genes (Modrek et al., 2001). Many different studies predict that from 40 to 60% of human genes are alternatively spliced (Croft et al., 2000; Hide et al., 2001; Kan et al., 2001; Modrek et al., 2001). Alternative splicing occurs in *C. elegans* and plays an important role in development. This chapter builds upon a previous review in *C. elegans* II of constitutive splicing signals in *C. elegans* (Blumenthal and Steward, 1997).



**Figure 1. Types of alternative splicing.** In these graphics, exons are represented by boxes and introns by lines. Exon regions included in the messages by alternative splicing are colored while constitutive exons are shown in gray. Promoters are indicated with arrows and polyadenylation sites with AAAA.

Many genes are alternatively spliced in tissue-specific, developmentally-regulated, and hormone-responsive manners, providing an additional mechanism for regulation of gene expression (Bandman, 1992; Caceres and Kornblihtt, 2002; Grabowski, 1998; Grabowski and Black, 2001; Singh, 2002; Stamm et al., 1994). Types of alternative splicing include the use of alternative 5' splice sites, alternative 3' splice sites, cassette exons, retained introns and mutually exclusive exons (Figure 1). Alternative splicing often leads to changes in the primary amino acid sequence of the protein, sometimes subtle and sometimes quite dramatic. Alternative splicing can control where the site of 3' end cleavage and polyadenylation occurs, indicating a strong interaction between the splicing and 3' end formation machineries. Alternative splicing can lead to messages with premature termination codons, for example by including a cassette exon with an in-frame stop codon. These premature termination codon-containing messages can be substrates for the nonsense-mediated decay (NMD) pathway in *C. elegans* (Mango, 2001). Evidence for this was first demonstrated in *C. elegans* in a genetic screen for alternative splicing factors. Morrison et al. showed that two of the genes encoding SR protein splicing factors have alternative isoforms with premature termination codons. These alternatively spliced forms were preferentially stabilized in suppressor with morphogenetic effects on genitalia (*smg*) mutants (Morrison et al., 1997). This use of alternative splicing to generate NMD substrates is very common in mammals (Lewis et al., 2003). In humans, polypyrimidine tract binding protein (PTB), which has multiple roles in mRNA processing including alternative splicing regulation, 3' end formation, and stability, regulates its own abundance by influencing the splicing of its own message. In this negative feedback loop, PTB promotes skipping of its own exon 11, leading to an isoform with an in-frame stop codon that is a substrate for NMD (Wollerton et al., 2004). In addition to turning off transcription to downregulate gene expression, alternative splicing is commonly used to downregulate gene expression by producing messages with truncated open reading frames that are rapidly degraded by NMD.

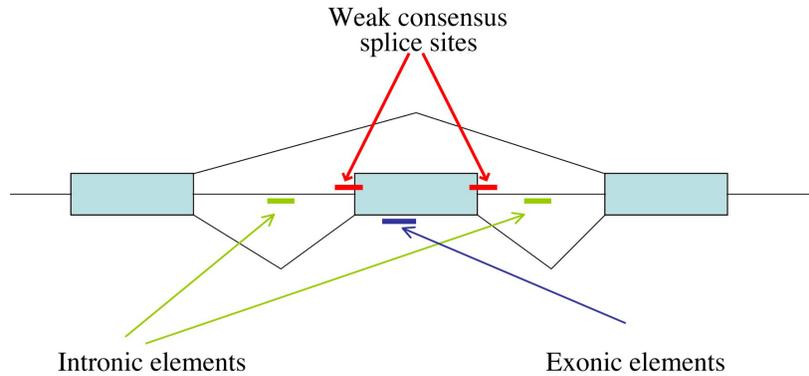
## 2. Splicing of precursors to messenger RNA

Alternative splicing is regulated at the sites on the pre-mRNA where the splicing machinery assembles. The two transesterification reaction steps of pre-mRNA splicing occur in a large multi-component ribonucleoprotein complex, the spliceosome (Burge et al., 1999; Caceres and Krainer, 1997; Kramer, 1996). The spliceosome assembles onto the pre-mRNA via an ordered binding of its subunits, the uridine-rich small nuclear ribonucleoprotein complexes (U snRNPs). SnRNP assembly is guided by RNA-RNA interactions between the U snRNAs and the pre-mRNA and the snRNAs with each other. First U1 binds to the pre-mRNA guided by base-pairing interactions between its 5' end and the 5' splice site. U2 binding is guided by base pairing interactions with the branchpoint sequence in the intron. Then the U4/U6-U5 tri-snRNP is recruited to the forming spliceosome. Extensive rearrangements of RNA-RNA interactions then occur in the spliceosome (Ares and Weiser, 1994; Madhani and Guthrie, 1994). There is evidence from recent studies that the spliceosomal components can be found together in a pre-assembled penta-snRNP complex (Stevens et al., 2002). The mammalian penta-snRNP has been shown to assemble onto 5' splice sites and undergo rearrangements that replace U1 interactions with U5 interactions (Malca et al., 2003). The functional importance of the pre-assembled penta-snRNP compared to models of step-wise spliceosome assembly on the substrate pre-mRNA is still being studied.

Several classes of splicing factor proteins have been identified as essential for metazoan spliceosome assembly yet are not integral components of the snRNPs (Kramer, 1996). The SR protein family of splicing factors is required for the earliest interactions of U1 snRNP with the pre-mRNA and for subsequent spliceosome assembly steps (Fu, 1995; Graveley, 2000; Manley and Tacke, 1996). SR proteins have redundant functions in promoting constitutive splicing, but certain family members have distinct roles in promoting specific alternative splice site usage. SR protein family members have distinct binding specificities for sequences within exons that serve as splicing enhancer elements (Cartegni and Krainer, 2002). U2 auxiliary factor (U2AF) is a heterodimeric splicing factor composed of 35 and 65 kD subunits. It binds to the polypyrimidine tract and recognizes the AG dinucleotide at the 3' end of the intron to promote U2 snRNA interactions with the intron branchpoint sequence (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999b). U2AF, SR proteins and U1 snRNP assemble onto the pre-mRNA in an ATP-independent manner to form the early (E) complex (Michaud and Reed, 1991; Staknis and Reed, 1994). It is important to note that assembly of the E complex essentially identifies both the 5' end 3' splice site for further spliceosome assembly, and thus control of this assembly can be a way of regulating alternative splice site choice. Splicing factors interact with the pre-mRNA, with other splicing factors and with protein components of the snRNPs (Burge et al., 1999; Caceres and Krainer, 1997; Kramer, 1996). This cooperative assembly increases the local concentration of the snRNAs in the vicinity of splice sites and therefore is likely to promote the RNA-RNA interactions required for spliceosome assembly.

## 3. Models for alternative splicing regulation

From research over the past 20 years, some general themes have emerged for alternative splicing regulation, although the exact mechanisms still need to be determined. Alternatively spliced exons often have weak consensus sequences at the 5' and 3' ends of the introns, suggesting that additional signals are required for recognition of the exon by the splicing machinery (Lopez, 1998). *Cis*-acting pre-mRNA sequences responsible for regulation of splicing have been identified for many genes (Cooper and Mattox, 1997). These regions are found in exons or in introns and can be enhancers or silencers of splice site usage (Figure 2). These sequence motifs serve as binding sites for protein factors that can enhance or inhibit the ability of the spliceosome to recognize the exons. The exonic elements not only encode amino acids but also regulate their own ability to be spliced into the mature message. *Trans*-acting splicing factors that interact with splicing regulatory elements in exons have been identified. Subsets of the SR proteins bind with regulatory sequences important for splicing control (Gontarek and Derse, 1996; Kanopka et al., 1996; Lavigne et al., 1993; Nagel et al., 1998; Ramchatesingh et al., 1995; Staknis and Reed, 1994; Sun et al., 1993). Heterogeneous nuclear ribonucleoprotein (hnRNP) A/B family members can bind to high-affinity sequences in exons and inhibit splicing through blocking SR proteins from binding to the exons (Caputi et al., 1999; Zhu et al., 2001).



**Figure 2. Locations of regions on the pre-mRNA that can affect alternative splicing.** Some combination of these regulatory regions can usually be found. Weaker consensus splice sites surrounding the alternative exon, exonic regulatory regions and intronic regulatory regions are indicated.

Splicing factors important for tissue-specific regulation of vertebrate splicing often assemble into multicomponent complexes on intronic splicing regulatory elements. The downstream control sequence (DCS), found in the intron downstream of the human neural-specific c-src N1 exon, is one such example. Factors that bind to the DCS regulate N1 splicing; these include hnRNP H, hnRNP F, KH-type splicing regulatory protein (KSRP) and a neural-specific homolog of PTB (nPTB; Chou et al., 1999; Markovtsov et al., 2000; Min et al., 1995; Min et al., 1997). Comparison of the factors that assemble onto this element from neuronal cell nuclear extract vs. epithelial cell nuclear extract indicate that a subset of these proteins bind from both extracts. Factors in the neuronal extract promote assembly of a different complex that is required to activate splicing of the neural-specific exon. PTB was identified by its ability to bind to polypyrimidine tracts. It has been shown to regulate the alternative splicing of its own message, as well as others including cardiac troponin T, alpha-actinin, fibroblast growth factor receptor R2, calcitonin/CGRP, and alpha tropomyosin pre-mRNAs (Wagner and Garcia-Blanco, 2001; Wollerton et al., 2001). In these systems, PTB binding to both introns flanking an exon promotes exon skipping (Wagner and Garcia-Blanco, 2001). CUG binding protein and related family members known as CELF proteins interact with CUG repeats in introns to regulate splicing (Ladd et al., 2001; Philips et al., 1998). An antagonistic interaction between one of the CELF proteins, ETR-3, and PTB regulates troponin t alternative splicing (Charlet-B et al., 2002). The nova-1 protein is important for regulation of alternative splicing in the nervous system. This neuronal protein binds the sequence UCAY (Jensen, Dredge, et al., 2000; Jensen, Musunuru, et al., 2000). A balance between nova activity and PTB is important in regulation of alternative splicing in neurons (Polydorides et al., 2000). The mechanism by which binding of these factors to intronic elements regulates exon inclusion is not clear. Genomic analysis indicates that homologs of all of the major vertebrate splicing factors discussed above can also be found in *C. elegans* (AMZ, unpublished observations).

#### 4. Constitutive splicing factors in *C. elegans*

The spliceosomal snRNAs of *C. elegans* are very similar to their yeast and mammalian counterparts (Thomas et al., 1990). The importance of U1 snRNA base-pairing with the 5' splice site is demonstrated by two extragenic allele-specific suppressors of mutations of the canonical G that begins *C. elegans* introns. *sup-6* and *sup-39* encode two of the 12 U1 snRNA genes in *C. elegans* and were identified in screens as suppressors that can allow mutated splice donors to be recognized by the splicing machinery (Roller et al., 2000; Run et al., 1996). The suppressor mutations are compensatory mutations in the 5' end of U1 snRNA genes (Zahler et al., 2004). These U1 snRNA suppressors are informational suppressors, allele-specific gene-nonspecific suppressors, and join the amber suppressors and the *smg* genes as members of this group (Mount and Anderson, 2000).

*C. elegans* contains seven different members of the SR protein family called the *rsp* genes (Kawano et al., 2000; Longman et al., 2000; Morrison et al., 1997). Family members are expressed in all cells at all stages of development (Kawano et al., 2000). There is functional overlap among the family members as studies suggest that, with the exception of *rsp-3*, multiple SR proteins need to be targeted by RNA interference in order for a phenotype to become evident (Kawano et al., 2000; Longman et al., 2000; Longman et al., 2001). The phenomenon of trans splicing is common in nematode species (see Trans-splicing and operons). Using splicing competent embryonic extracts from parasitic nematode *Ascaris lumbricoides*, Sanford and Bruzik have shown that SR proteins are essential for the trans splicing reaction in nematodes and that the phosphorylation state and activity level of these proteins change during early development (Sanford and Bruzik, 1999a; Sanford and Bruzik, 1999b).

*C. elegans* has a homolog of each U2AF subunit, *uaf-1* and *uaf-2* (Zorio and Blumenthal, 1999a; Zorio et al., 1997). Using the short defined polypyrimidine tract at the 3' end of *C. elegans* introns (Blumenthal and Steward, 1997), studies in this organism were able to demonstrate the specificity with which the U2AF heterodimer subunits recognize the polypyrimidine tract and 3' splice site at the end of the intron. The U2AF65 homolog, *uaf-1*, interacts with the polypyrimidine tract while the U2AF35 homolog *uaf-2* recognizes the AG dinucleotide at the 3' end of the intron (Zorio and Blumenthal, 1999b). This showed that even though the AG at the 3' end of the intron is involved in the second transesterification reaction of splicing, it also has an important role in early spliceosome assembly. *UAF-1* regulates its own expression. The *uaf-1* message is alternatively spliced to contain an unusual exon with an in-frame stop codon that should be a subject for nonsense mediated decay. This exon is unusual because it contains 10 matches to the *C. elegans* 3' splice acceptor sequence UUUUCAG/R, the sequence recognized directly by *C. elegans* U2AF (Zorio and Blumenthal, 1999b; Zorio et al., 1997). MacMorris et al. showed that these splice acceptor repeats cause the pre-mRNA to be retained in the nucleus and thus not become a substrate for translation or nonsense-mediated decay. Experiments in which the UUUUCAG/R repeats were placed in a GFP transgene indicate that *UAF-1* is responsible for the nuclear retention of its message (MacMorris et al., 1999).

## 5. Identification of alternatively spliced genes in *C. elegans*

Intron and exon structure in *C. elegans* is similar to that found in other higher eukaryotes (Blumenthal and Steward, 1997). The completed genome sequence and >200,000 expressed sequence tag (EST) complementary DNA (cDNA) sequences are available in public databases such as Wormbase at <http://www.wormbase.org/> (*C. elegans* Sequencing Consortium, 1998; Stein et al., 2001). Computational approaches have been developed to identify alternatively spliced *C. elegans* genes by comparing cDNA and genomic sequences (Kent and Zahler, 2000). This led to the initial identification of 844 alternatively spliced genes. Regularly updated lists of alternatively spliced genes can be found at the Intronerator web site, <http://genome-test.cse.ucsc.edu/Intronerator/>. Wormbase annotations of genes indicate alternatively spliced isoforms. These are generated by hand annotation based on cDNA and EST data as well as experimental data reported by researchers. As of release WS132, there are 2,562 alternatively spliced genes annotated in Wormbase, accounting for 13% of annotated *C. elegans* genes. This is likely an underestimate of the total percentage of *C. elegans* genes that are alternatively spliced as these values are dependent on EST and cDNA coverage of the genome and at the current time 25% of the 19,726 annotated *C. elegans* coding sequences have no cDNA transcriptional evidence (Wormbase WS132 release notes). Genes with alternative isoforms are denoted by a lower case letter after the assigned gene name, for example T01D7.1a. Many of the genes with alternative isoforms are generated by using alternative promoters with unique first exons that splice into a common second exon. Even though these alternative promoters will lead to alternative 5' exons spliced to common downstream exons (Figure 1), the generation of the unique first exon is not regulated at the level of alternative splicing but instead by transcription.

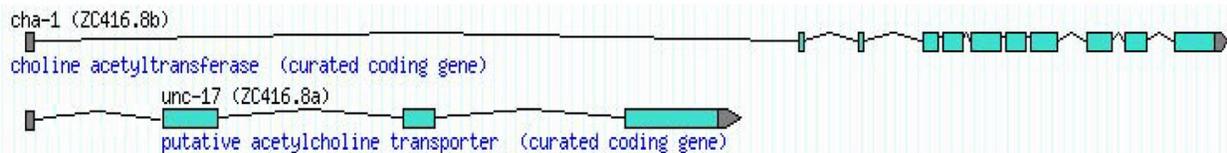
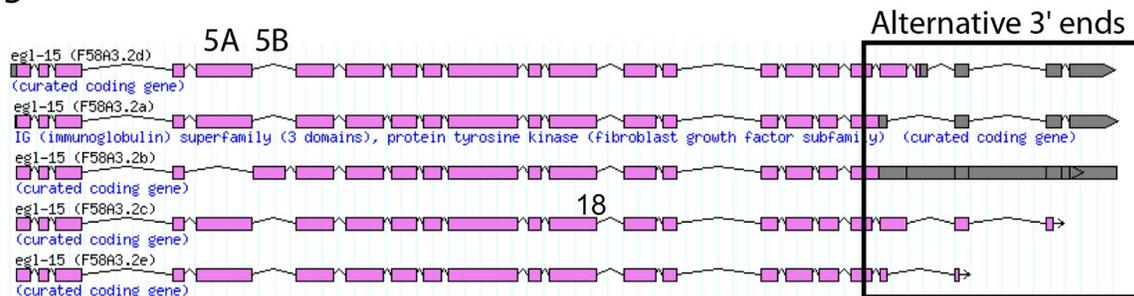
## 6. Some examples of alternatively spliced *C. elegans* genes

There are four genes in *C. elegans* that encode the  $\alpha$  subunit of the vacuolar ATPase, suggesting a mechanism for regulating the activity of this pump (Oka et al., 2001). One of these genes, encoded by *unc-32*, can be alternatively spliced to yield six different isoforms, adding even more complexity (Figure 3A). There are three different mutually exclusive exons for exon 4, and two for exon 7 leading to six possible transcripts (Pujol et al., 2001). In the study by Pujol et al., GFP was fused in frame to exon 4B of an *unc-32* genomic clone and transfected into animals. The expression of GFP was limited to neurons suggesting that splicing of exon 4B is neural-specific (Pujol et al., 2001).

Alternative splicing can be used to transcribe two distinct genes from a single promoter, in a way that accomplishes the same end but is distinct from trans-splicing of operons. This occurs for the *cha-1* and *unc-17* genes. *cha-1* encodes choline acetyl-transferase while *unc-17* encodes a synaptic vesicle-associated acetylcholine transporter. Analysis of transcripts of these two genes involved in acetylcholine metabolism indicate that they share a common promoter and 5' untranslated exon (Figure 3B). The *unc-17* transcript is nested within the long first intron of *cha-1* (Alfonso et al., 1994). In this case, two proteins with related functions but with no peptide sequences in common are produced as a result of alternative splicing of a common mRNA precursor.

The fibroblast growth factor receptor homolog *egl-15* is alternatively spliced to yield five different C-termini. In addition, the gene has two distinct mutually exclusive central exons, 5A and 5B (Figure 3C). These two exons encode a domain in the extracellular portion of the protein that gives ligand specificity to the receptor. Exon 5A is required for the response to the FGF chemoattractant *EGL-17*, which guides migrating sex myoblasts to their final

positions. Exon 5B is required for the essential functions mediated by this receptor through interaction with the ligand LET-756 (Goodman et al., 2003). In this case of alternative splicing, two mutually exclusive exons allow one gene to code for a receptor tyrosine kinase with two distinct functional specificities.

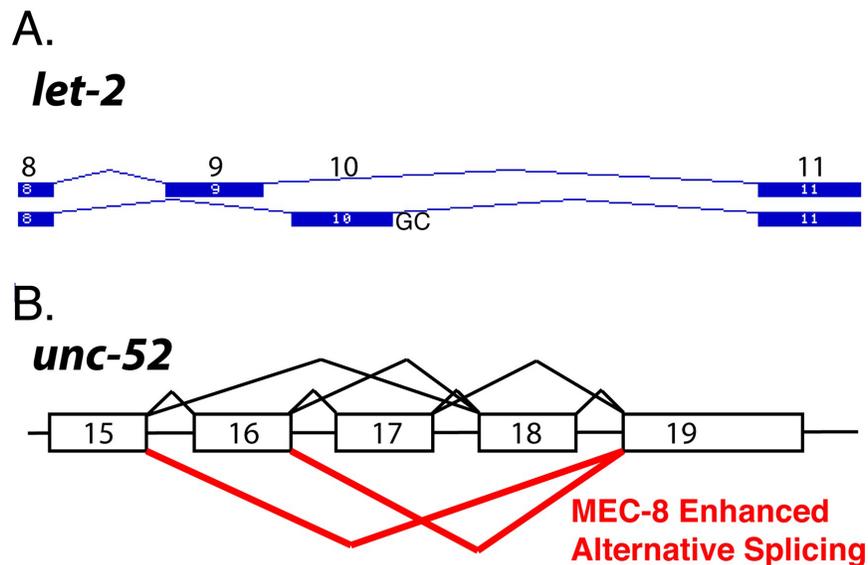
**A.***unc-32***B.***cha-1/unc-17***C.***egl-15*

**Figure 3. Examples of *C. elegans* alternatively spliced genes.** Images are screen shots taken from the Wormbase genome browser. (A) *unc-32*. The three alternative exon 4s and two alternative exon 7s are indicated. (B) *cha-1* and *unc-17* share a common promoter and 5' untranslated exon. (C) *egl-15*. The two mutually exclusive exons, 5A and 5B, which provide functional specificity, are indicated, as are the five different 3' ends generated by alternative splicing. Both exon 5 isoforms have been detected with each of the 3' end isoforms, so there are 10 alternatively spliced isoforms of this gene (Goodman et al., 2003). Only five of the isoforms are indicated in the graphic.

## 7. Regulation of alternative splicing in *C. elegans*

Splicing studies in *C. elegans* have demonstrated the importance of varying strengths among splice site consensus sequences to allow for splicing regulation. 5' splice site strength is determined by the extent of complementarity of the splice donor region to the 5' end of U1 snRNA. The alpha 2(IV) collagen gene *let-2* is alternatively spliced to yield two forms of the protein containing one of two mutually exclusive exons, exon 9 or exon 10 (Figure 4A). There is developmental regulation of this alternative splicing; 95% of embryonic LET-2

messages contain exon 9 while in adults 90% contain exon 10. The larval stages show a gradual shift in usage between these isoforms, indicating developmental control of alternative splicing (Sibley et al., 1993). The intron downstream of *let-2* exon 10 begins with a rare GC dinucleotide instead of the canonical GU. Only 0.5% of introns in *C. elegans* begin with a GC dinucleotide (Farrer et al., 2002). In *Ascaris*, which maintains the same developmentally regulated alternative splicing of this alpha 2 (IV) collagen gene, the same intron begins with GU (Pettitt and Kingston, 1994). The GC splice donor serves as a weak consensus sequence that allows for the proper regulation of alternative splicing. Replacement of this sequence with a GU donor (such as the one from *Ascaris*) that has weak overall base-pairing with the 5' end of U1 snRNA, allows for maintenance of splicing regulation. Replacing the GC donor with a moderate or strong consensus GU donor abolishes the splicing regulation and leads to high usage of exon 10 in embryos (Farrer et al., 2002). The presence of a weak splice donor for exon 10 allows developmentally regulated factors to control the recognition of this exon by the splicing machinery. The identity of these factors is not yet known, but our group has identified an 11 base sequence element in the intron between exons 10 and 11 that is required for efficient inclusion of exon 10 in late larval stages. This element may serve as an interacting site for developmentally regulated splicing factors (Tracy Farrer and AMZ, unpublished observation).



**Figure 4. Alternative splicing regulation in *C. elegans* genes.** (A) Alternative splicing of the *let-2* gene containing mutually exclusive exons 9 and 10. The intron downstream of exon 10 begins with a GC dinucleotide important for splicing regulation. (B) Diagram of alternative splicing of the *unc-52* gene. Every combination of exon inclusion and skipping of exons 16, 17 and 18 has been detected. MEC-8 enhances the splicing of two of the isoforms, 15-19 and 15-16-19, indicated in red (Lundquist et al., 1996).

Genetic analysis in *C. elegans* has led to the identification of splicing regulatory pathways. MEC-8 is a two RNA recognition motif (RRM)-containing, nuclear protein. Mutations in *mec-8* lead to mechanosensory and other defects. Mutations in *mec-8* are synthetically lethal with viable mutations in the *unc-52* gene (Lundquist et al., 1996). UNC-52 is a homolog of the mammalian perlecan protein, an extracellular matrix molecule important in muscle anchorage (Rogalski et al., 1993). These viable *unc-52* mutations occur in the alternatively spliced exons 17 and 18 (Figure 4b). Every possible combination of inclusion or skipping of exons 16, 17 and 18 has been detected (Rogalski et al., 1995). Skipping of exons 17 and 18 is dependent on the function of *mec-8* (Lundquist et al., 1996). MEC-8 protein is found in all nuclei in embryos, but only in a subset of mechanosensory neurons and hypodermal cells in adults (Spike et al., 2002). *unc-52* mutant alleles with stop codons in exons 17 and 18 develop a progressive paralysis at later developmental stages. Null mutations in the *unc-52* gene lead to embryonic lethality (Rogalski et al., 1995). Animals containing both viable mutations in *unc-52* and viable mutations in *mec-8* have a synthetic lethal phenotype that mimics the embryonic lethality of the lethal *unc-52* alleles (Lundquist and Herman, 1994). In *mec-8(+)* animals with viable *unc-52* mutations, there is a loss of full-length UNC-52 production in the later developmental stages, leading to the adult paralysis phenotype. In *mec-8; unc-52* double mutant animals, loss of *mec-8*-enhanced skipping of exons 17 and 18 (containing stop codons) leads to insufficient production of full-length UNC-52 in embryos and thus to the more severe embryonic lethality. Two suppressors of the *mec-8; unc-52* synthetic lethality were identified (Lundquist and Herman, 1994). One of these, *smu-1*, which encodes a nuclear protein that contains five WD motifs, is ubiquitously expressed and is 62% identical to a human spliceosome associated protein fSAP57 (Spike et al., 2001). *smu-2* encodes a *C. elegans* homolog of the RED protein, which has also been shown to be associated with the human spliceosome. SMU-1 and SMU-2 bind to each other; SMU-2

stabilizes SMU-1 but SMU-1 does not stabilize SMU-2. Based on their homology with components of the mammalian spliceosome, it is hypothesized that SMU-1 and SMU-2 might interact with the spliceosome and modulate splice site choice (Spartz et al., 2004).

Alternative splicing factors play a role in *C. elegans* sex determination. In hermaphrodites, the feminizing locus on X (*fox-1*) gene is a regulator of sex determination that acts as a numerator for counting X chromosomes (Hodgkin et al., 1994). *fox-1* encodes an RRM protein that acts post-transcriptionally to inhibit the expression of the *xol-1* gene, the major specifier of male fate. FOX-1 inhibits splicing of the terminal intron of *xol-1* pre-mRNA, preventing production of active XOL-1 (Nicoll et al., 1997).

Tissue specific expression of alternative splicing factors has been demonstrated in *C. elegans*. Two genetically identified neuronal, nuclear RNA binding proteins involved in synaptic transmission have been cloned in *C. elegans*. These are *unc-75*, which is an ortholog of the human CELF family of alternative splicing factors, and *exc-7*, which is a homolog of the *Drosophila* neuronal splicing factor ELAV (Loria et al., 2003). Another ELAV homolog in *C. elegans*, *etr-1*, is expressed only in muscle cells and is essential for development (Milne and Hodgkin, 1999).

## 8. Conclusions

With its many alternatively spliced genes and evolutionarily conserved alternative splicing factors, *C. elegans* offers a tractable model system for the study of pre-mRNA splicing regulation. With much smaller average intron size compared to vertebrates, a fully sequenced genome, and a highly developed and simple genetics system, research in *C. elegans* has the potential to provide important contributions to our understanding of alternative splicing regulation.

## 9. Acknowledgements

I am grateful to the members of my laboratory at UC Santa Cruz for many helpful comments on this manuscript. Research in my laboratory is supported by grant GM61646 from the National Institute of General Medical Sciences.

## 10. References

- Alfonso, A., Grundahl, K., McManus, J.R., Asbury, J.M., and Rand, J.B. (1994). Alternative splicing leads to two cholinergic proteins in *Caenorhabditis elegans*. *J. Mol. Biol.* 241, 627–630. [Abstract Article](#)
- Ares, M., Jr., and Weiser, B. (1994). Rearrangement of snRNA structure during assembly and function of the spliceosome. *Prog. Nucleic Acid. Res. Mol. Biol.* 50, 131–159. [Abstract](#)
- Bandman, E. (1992). Contractile protein isoforms in muscle development. *Dev. Biol.* 154, 273–283. [Abstract Article](#)
- Blumenthal, T., and Steward, K. (1997). RNA processing and gene structure. In *C. elegans II*, D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, eds. (Plainview, NY, Cold Spring Harbor Laboratory Press), pp. 117–145.
- Burge, C.B., Tuschl, T., and Sharp, P.A. (1999). Splicing of precursors to mRNAs by the spliceosome. In *The RNA World*, 2nd Edition, R.F. Gesteland, T.R. Cech, and J.F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 525–560.
- C. elegans* Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018. [Abstract Article](#)
- Caceres, J.F., and Kornblihtt, A.R. (2002). Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.* 18, 186–193. [Abstract Article](#)
- Caceres, J.F., and Krainer, A.R. (1997). Mammalian pre-mRNA splicing factors. In *Eukaryotic mRNA Processing.*, A.R. Krainer, ed. (Oxford, UK: IRL Press), pp. 174–212.
- Caputi, M., Mayeda, A., Krainer, A.R., and Zahler, A.M. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J.* 18, 4060–4067. [Abstract Article](#)

- Cartegni, L., and Krainer, A.R. (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat. Genet.* *30*, 377–384. [Abstract Article](#)
- Charlet-B, N., Logan, P., Singh, G., and Cooper, T.A. (2002). Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing. *Mol. Cell* *9*, 649–658. [Abstract Article](#)
- Chou, M.Y., Rooke, N., Turck, C.W., and Black, D.L. (1999). hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Mol. Cell. Biol.* *19*, 69–77. [Abstract](#)
- Cooper, T.A., and Mattox, W. (1997). The regulation of splice-site selection, and its role in human disease. *Am. J. Hum. Genet.* *61*, 259–266. [Abstract](#)
- Croft, L., Schandorff, S., Clark, F., Burrage, K., Arctander, P., and Mattick, J.S. (2000). ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome. *Nat. Genet.* *24*, 340–341. [Abstract Article](#)
- Farrer, T., Roller, A.B., Kent, W.J., and Zahler, A.M. (2002). Analysis of the role of *Caenorhabditis elegans* GC-AG introns in regulated splicing. *Nucleic Acids Res.* *30*, 3360–3367. [Abstract Article](#)
- Fu, X.D. (1995). The superfamily of arginine/serine-rich splicing factors. *RNA* *1*, 663–680. [Abstract](#)
- Gontarek, R.R., and Derse, D. (1996). Interactions among SR proteins, an exonic splicing enhancer, and a lentivirus rev protein regulate alternative splicing. *Mol. Cell. Biol.* *16*, 2325–2331. [Abstract](#)
- Goodman, S.J., Branda, C.S., Robinson, M.K., Burdine, R.D., and Stern, M.J. (2003). Alternative splicing affecting a novel domain in the *C. elegans* EGL-15 FGF receptor confers functional specificity. *Development* *130*, 3757–3766. [Abstract Article](#)
- Grabowski, P.J. (1998). Splicing regulation in neurons: tinkering with cell-specific control. *Cell* *92*, 709–712. [Abstract Article](#)
- Grabowski, P.J., and Black, D.L. (2001). Alternative RNA splicing in the nervous system. *Prog Neurobiol* *65*, 289–308. [Abstract Article](#)
- Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. *RNA* *6*, 1197–1211. [Abstract Article](#)
- Hide, W.A., Babenko, V.N., van Heusden, P.A., Seoighe, C., and Kelso, J.F. (2001). The contribution of exon-skipping events on chromosome 22 to protein coding diversity. *Genome Res.* *11*, 1848–1853. [Abstract Article](#)
- Hodgkin, J., Zellan, J.D., and Albertson, D.G. (1994). Identification of a candidate primary sex determination locus, *fox-1*, on the X chromosome of *Caenorhabditis elegans*. *Development* *120*, 3681–3689. [Abstract](#)
- International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* *431*, 931–945. [Abstract Article](#)
- Jensen, K.B., Dredge, B.K., Stefani, G., Zhong, R., Buckanovich, R.J., Okano, H.J., Yang, Y.Y., and Darnell, R.B. (2000). Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron* *25*, 359–371. [Abstract Article](#)
- Jensen, K.B., Musunuru, K., Lewis, H.A., Burley, S.K., and Darnell, R.B. (2000). The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proc. Natl. Acad. Sci. USA* *97*, 5740–5745. [Abstract Article](#)
- Kan, Z., Rouchka, E.C., Gish, W.R., and States, D.J. (2001). Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res.* *11*, 889–900. [Abstract Article](#)
- Kanopka, A., Muhlemann, O., and Akusjarvi, G. (1996). Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* *381*, 535–538. [Abstract Article](#)

- Kawano, T., Fujita, M., and Sakamoto, H. (2000). Unique and redundant functions of SR proteins, a conserved family of splicing factors, in *Caenorhabditis elegans* development. *Mech. Dev.* *95*, 67–76. [Abstract Article](#)
- Kent, W.J., and Zahler, A.M. (2000). The intronerator: exploring introns and alternative splicing in *Caenorhabditis elegans*. *Nucleic Acids Res.* *28*, 91–93. [Abstract Article](#)
- Kramer, A. (1996). The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem* *65*, 367–409. [Abstract Article](#)
- Ladd, A.N., Charlet, N., and Cooper, T.A. (2001). The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol. Cell. Biol.* *21*, 1285–1296. [Abstract Article](#)
- Lavigueur, A., La Branche, H., Kornblihtt, A.R., and Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes Dev.* *7*, 2405–2417. [Abstract](#)
- Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. USA* *100*, 189–192. [Abstract Article](#)
- Longman, D., Johnstone, I.L., and Caceres, J.F. (2000). Functional characterization of SR and SR-related genes in *Caenorhabditis elegans*. *EMBO J.* *19*, 1625–1637. [Abstract Article](#)
- Longman, D., McGarvey, T., McCracken, S., Johnstone, I.L., Blencowe, B.J., and Caceres, J.F. (2001). Multiple interactions between SRm160 and SR family proteins in enhancer-dependent splicing and development of *C. elegans*. *Curr. Biol.* *11*, 1923–1933. [Abstract Article](#)
- Lopez, A.J. (1998). Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu. Rev. Genet.* *32*, 279–305. [Abstract Article](#)
- Loria, P.M., Duke, A., Rand, J.B., and Hobert, O. (2003). Two neuronal, nuclear-localized RNA binding proteins involved in synaptic transmission. *Curr. Biol.* *13*, 1317–1323. [Abstract Article](#)
- Lundquist, E.A., and Herman, R.K. (1994). The *mec-8* gene of *Caenorhabditis elegans* affects muscle and sensory neuron function and interacts with three other genes: *unc-52*, *smu-1* and *smu-2*. *Genetics* *138*, 83–101. [Abstract](#)
- Lundquist, E.A., Herman, R.K., Rogalski, T.M., Mullen, G.P., Moerman, D.G., and Shaw, J.E. (1996). The *mec-8* gene of *C. elegans* encodes a protein with two RNA recognition motifs and regulates alternative splicing of *unc-52* transcripts. *Development* *122*, 1601–1610. [Abstract](#)
- MacMorris, M.A., Zorio, D.A., and Blumenthal, T. (1999). An exon that prevents transport of a mature mRNA. *Proc. Natl. Acad. Sci. USA* *96*, 3813–3818. [Abstract Article](#)
- Madhani, H.D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* *28*, 1–26. [Abstract Article](#)
- Malca, H., Shomron, N., and Ast, G. (2003). The U1 snRNP base pairs with the 5' splice site within a penta-snRNP complex. *Mol. Cell. Biol.* *23*, 3442–3455. [Abstract Article](#)
- Mango, S.E. (2001). Stop making nonSense: the *C. elegans smg* genes. *Trends Genet.* *17*, 646–653. [Abstract Article](#)
- Manley, J.L., and Tacke, R. (1996). SR proteins and splicing control. *Genes Dev.* *10*, 1569–1579. [Abstract](#)
- Markovtsov, V., Nikolic, J.M., Goldman, J.A., Turck, C.W., Chou, M.Y., and Black, D.L. (2000). Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. *Mol. Cell. Biol.* *20*, 7463–7479. [Abstract Article](#)
- Merendino, L., Guth, S., Bilbao, D., Martinez, C., and Valcarcel, J. (1999). Inhibition of *msl-2* splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* *402*, 838–841. [Abstract Article](#)

- Michaud, S., and Reed, R. (1991). An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev.* *5*, 2534–2546. [Abstract](#)
- Milne, C.A., and Hodgkin, J. (1999). ETR-1, a homologue of a protein linked to myotonic dystrophy, is essential for muscle development in *Caenorhabditis elegans*. *Curr. Biol.* *9*, 1243–1246. [Abstract Article](#)
- Min, H., Chan, R.C., and Black, D.L. (1995). The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. *Genes Dev.* *9*, 2659–2671. [Abstract](#)
- Min, H., Turck, C.W., Nikolic, J.M., and Black, D.L. (1997). A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes Dev.* *11*, 1023–1036. [Abstract](#)
- Modrek, B., Resch, A., Grasso, C., and Lee, C. (2001). Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res.* *29*, 2850–2859. [Abstract Article](#)
- Morrison, M., Harris, K.S., and Roth, M.B. (1997). *smg* mutants affect the expression of alternatively spliced SR protein mRNAs in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* *94*, 9782–9785. [Abstract Article](#)
- Mount, S.M., and Anderson, P. (2000). Expanding the definition of informational suppression. *Trends Genet.* *16*, 157. [Abstract Article](#)
- Nagel, R.J., Lancaster, A.M., and Zahler, A.M. (1998). Specific binding of an exonic splicing enhancer by the pre-mRNA splicing factor SRp55. *RNA* *4*, 11–23. [Abstract](#)
- Nicoll, M., Akerib, C.C., and Meyer, B.J. (1997). X-chromosome-counting mechanisms that determine nematode sex. *Nature* *388*, 200–204. [Abstract Article](#)
- Oka, T., Toyomura, T., Honjo, K., Wada, Y., and Futai, M. (2001). Four subunit isoforms of *Caenorhabditis elegans* vacuolar H<sup>+</sup>-ATPase. Cell-specific expression during development. *J. Biol. Chem.* *276*, 33079–33085. [Abstract Article](#)
- Pettitt, J., and Kingston, I.B. (1994). Developmentally regulated alternative splicing of a nematode type IV collagen gene. *Dev. Biol.* *161*, 22–29. [Abstract Article](#)
- Philips, A.V., Timchenko, L.T., and Cooper, T.A. (1998). Disruption of Splicing Regulated by a CUG-Binding Protein in Myotonic Dystrophy. *Science* *280*, 737–741. [Abstract Article](#)
- Polydorides, A.D., Okano, H.J., Yang, Y.Y., Stefani, G., and Darnell, R.B. (2000). A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc. Natl. Acad. Sci. USA* *97*, 6350–6355. [Abstract Article](#)
- Pujol, N., Bonnerot, C., Ewbank, J.J., Kohara, Y., and Thierry-Mieg, D. (2001). The *Caenorhabditis elegans unc-32* gene encodes alternative forms of a vacuolar ATPase subunit. *J. Biol. Chem.* *276*, 11913–11921. [Abstract Article](#)
- Ramchatesingh, J., Zahler, A.M., Neugebauer, K.M., Roth, M.B., and Cooper, T.A. (1995). A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.* *15*, 4898–4907. [Abstract](#)
- Rogalski, T.M., Williams, B.D., Mullen, G.P., and Moerman, D.G. (1993). Products of the *unc-52* gene in *Caenorhabditis elegans* are homologous to the core protein of the mammalian basement membrane heparan sulfate proteoglycan. *Genes Dev.* *7*, 1471–1484. [Abstract](#)
- Rogalski, T.M., Gilchrist, E.J., Mullen, G.P., and Moerman, D.G. (1995). Mutations in the *unc-52* gene responsible for body wall muscle defects in adult *Caenorhabditis elegans* are located in alternatively spliced exons. *Genetics* *139*, 159–169. [Abstract](#)
- Roller, A.B., Hoffman, D.C., and Zahler, A.M. (2000). The allele-specific suppressor *sup-39* alters use of cryptic splice sites in *C. elegans*. *Genetics* *154*, 1169–1179. [Abstract](#)

- Run, J.Q., Steven, R., Hung, M.S., van Weeghel, R., Culotti, J.G., and Way, J.C. (1996). Suppressors of the *unc-73* gene of *Caenorhabditis elegans*. *Genetics* *143*, 225–236. [Abstract](#)
- Sanford, J.R., and Bruzik, J.P. (1999a). Developmental regulation of SR protein phosphorylation and activity. *Genes Dev.* *13*, 1513–1518. [Abstract](#)
- Sanford, J.R., and Bruzik, J.P. (1999b). SR proteins are required for nematode trans-splicing *in vitro*. *RNA* *5*, 918–928. [Abstract Article](#)
- Sibley, M.H., Johnson, J.J., Mello, C.C., and Kramer, J.M. (1993). Genetic identification, sequence, and alternative splicing of the *Caenorhabditis elegans* alpha 2(IV) collagen gene. *J. Cell Biol.* *123*, 255–264. [Abstract Article](#)
- Singh, R. (2002). RNA-protein interactions that regulate pre-mRNA splicing. *Gene Expr.* *10*, 79–92. [Abstract](#)
- Spartz, A.K., Herman, R.K., and Shaw, J.E. (2004). SMU-2 and SMU-1, *Caenorhabditis elegans* homologs of mammalian spliceosome-associated proteins RED and fSAP57, work together to affect splice site choice. *Mol. Cell Biol.* *24*, 6811–6823. [Abstract Article](#)
- Spike, C.A., Davies, A.G., Shaw, J.E., and Herman, R.K. (2002). MEC-8 regulates alternative splicing of *unc-52* transcripts in *C. elegans* hypodermal cells. *Development* *129*, 4999–5008. [Abstract Article](#)
- Spike, C.A., Shaw, J.E., and Herman, R.K. (2001). Analysis of *smu-1*, a gene that regulates the alternative splicing of *unc-52* pre-mRNA in *Caenorhabditis elegans*. *Mol. Cell Biol.* *21*, 4985–4995. [Abstract Article](#)
- Staknis, D., and Reed, R. (1994). SR proteins promote the first specific recognition of the pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol. Cell Biol.* *14*, 7670–7682. [Abstract](#)
- Stamm, S., Zhang, M.Q., Marr, T.G., and Helfman, D.M. (1994). A sequence compilation and comparison of exons that are alternatively spliced in neurons. *Nucleic Acids Res.* *22*, 1515–1526. [Abstract](#)
- Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J., and Spieth, J. (2001). WormBase: network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res.* *29*, 82–86. [Abstract Article](#)
- Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., and Abelson, J. (2002). Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell* *9*, 31–44. [Abstract Article](#)
- Sun, Q., Mayeda, A., Hampson, R.K., Krainer, A.R., and Rottman, F.R. (1993). General splicing factor ASF/SF2 promotes alternative splicing by binding to an exonic splicing enhancer. *Genes Dev.* *7*, 2598–2608. [Abstract](#)
- Thomas, J., Lea, K., Zucker-Aprison, E., and Blumenthal, T. (1990). The spliceosomal snRNAs of *Caenorhabditis elegans*. *Nucleic Acids Res.* *18*, 2633–2642. [Abstract](#)
- Wagner, E.J., and Garcia-Blanco, M.A. (2001). Polypyrimidine tract binding protein antagonizes exon definition. *Mol. Cell Biol.* *21*, 3281–3288. [Abstract Article](#)
- Wollerton, M.C., Gooding, C., Robinson, F., Brown, E.C., Jackson, R.J., and Smith, C.W. (2001). Differential alternative splicing activity of isoforms of polypyrimidine tract binding protein (PTB). *RNA* *7*, 819–832. [Abstract Article](#)
- Wollerton, M.C., Gooding, C., Wagner, E.J., Garcia-Blanco, M.A., and Smith, C.W. (2004). Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol. Cell* *13*, 91–100. [Abstract Article](#)
- Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* *402*, 832–835. [Abstract Article](#)

Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell* 8, 1351–1361. [Abstract Article](#)

Zorio, D.A., and Blumenthal, T. (1999a). U2AF35 is encoded by an essential gene clustered in an operon with RRM/cyclophilin in *Caenorhabditis elegans*. *RNA* 5, 487–494. [Abstract Article](#)

Zorio, D.A., and Blumenthal, T. (1999b). Both subunits of U2AF recognize the 3' splice site in *C. elegans*. *Nature* 402, 835–838. [Abstract Article](#)

Zorio, D.A., Lea, K., and Blumenthal, T. (1997). Cloning of *Caenorhabditis* U2AF65: an alternatively spliced RNA containing a novel exon. *Mol. Cell Biol.* 17, 946–953. [Abstract](#)

Zahler, A.M., Tuttle, J.D., and Chisholm, A.D. (2004). Genetic suppression of intronic +1G mutations by compensatory U1 snRNA changes in *Caenorhabditis elegans*. *Genetics* 167, 1689–1696. [Abstract Article](#)



All WormBook content, except where otherwise noted, is licensed under a [Creative Commons Attribution License](#).