High Qualitative and Quantitative Conservation of Alternative Splicing in Caenorhabditis elegans and Caenorhabditis briggsae

Jakob Lewin Rukov, Manuel Irimia, Søren Mørk, Viktor Karlovich Lund, Jeppe Vinther, and Peter Arctander

Molecular Evolution Group, Department of Molecular Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

Alternative splicing (AS) is an important contributor to proteome diversity and is regarded as an explanatory factor for the relatively low number of human genes compared with less complex animals. To assess the evolutionary conservation of AS and its developmental regulation, we have investigated the qualitative and quantitative expression of 21 orthologous alternative splice events through the development of 2 nematode species separated by 85–110 Myr of evolutionary time. We demonstrate that most of these alternative splice events present in Caenorhabditis elegans are conserved in Caenorhabditis briggsae. Moreover, we find that relative isoform expression levels vary significantly during development for 78% of the AS events and that this quantitative variation is highly conserved between the 2 species. Our results suggest that AS is generally tightly regulated through development and that the regulatory mechanisms controlling AS are to a large extent conserved during the evolution of Caenorhabditis. This strong conservation indicates that both major and minor splice forms have important functional roles and that the relative quantities in which they are expressed are crucial. Our results therefore suggest that the quantitative regulation of isoform expression levels is an intrinsic part of most AS events. Moreover, our results indicate that AS contributes little to transcript variation in Caenorhabditis genes and that gene duplication may be the major evolutionary mechanism for the origin of novel transcripts in these 2 species.

Introduction

Alternative splicing (AS) is widely regarded as an important contributor to transcriptome and proteome diversity (Maniatis and Tasic 2002; Modrek and Lee 2002). It has been estimated that 40–60% of all human genes (Mironov et al. 1999; Kan et al. 2001; Modrek et al. 2001; Modrek and Lee 2002) and 74% of multiexon human genes (Johnson et al. 2003) are alternatively spliced. As such, AS may be one explanation for the apparent lack of correlation between gene number and organism complexity (Kim et al. 2004; Suzuki and Hayashizaki 2004).

Few studies have investigated the evolutionary aspects of AS. Most of these are estimates of the amount of AS conserved between human and mouse based on computational analyses of expressed sequence tag (EST) databases and have primarily focused on exon skipping (reviewed in Lareau et al. 2004). In general, these computational studies show a relatively low level of conservation of AS. However, the variability of estimates is high, ranging from 11% (Yeo et al. 2005) to 83% (Nurtdinov et al. 2003) of events conserved in both species.

Splice variants can be separated into 2 forms: major forms, present in more than 50% of the transcripts and minor forms, present in less than 50%. According to computational studies, major form exons are the most conserved, with 61% (Kan et al. 2002) to 98% (Modrek and Lee 2003) of isoforms observed in both species. Minor forms are found to be mostly species specific, being conserved within a range of 8% (Kan et al. 2002) to 25% (Modrek and Lee 2003). Because evolutionary conservation is widely regarded as an indicator of function, the low conservation of the minor forms has cast doubt on the functional importance of these low-frequency isoforms (Modrek and Lee 2003). These have been suggested to be experimental noise of various methods (Modrek and Lee 2002) or biological noise due to aberrant splicing caused by spliceosomal errors (Sorek et al. 2004). Minor forms have also been suggested to be recent species-specific forms that are being tested under near-neutral fitness scenarios (Modrek and Lee 2003; Ast 2004; Krull et al. 2005).

However, single-gene studies have shown that both high- and low-frequency isoforms may be functional and expressed in a regulated manner (Morrison et al. 1997; Zorio et al. 1997; Park et al. 2003; Frey et al. 2005). Cases of tissue-specific expression (Mullen et al. 1999; Lazaridis et al. 2000; Tabuchi and Sudhoff 2002; Yeo et al. 2004), stage-specific expression (Mullen et al. 1999), and even subcellular compartmentalization (Laity et al. 2000; Tong et al. 2003) of the minor isoforms have been reported, supporting functionality for these isoforms. Accordingly, a recent definition of AS functionality has been suggested (Sorek et al. 2004): “an mRNA can be defined as being ‘functional’ if it is required during the life cycle of the organism and activated in a regulated manner.”

Furthermore, the importance of the quantitative aspects of AS is emerging. Various cases indicate that not only the presence of all isoforms is important but also the precise quantitative regulation of their expression. In the Piccolo protein C2A domain, AS of 9 amino acids changes the Ca$^{2+}$ sensitivity of the protein. This is speculated to fine-tune the Ca$^{2+}$ dependency of presynaptic exocytosis (Garcia et al. 2004). In Neurospora, the ratio of 2 splice variants of the frequency (frq) gene (long FRQ and short FRQ) is temperature dependent, and this stabilizes the circadian clock by increasing the temperature range over which it can function (Colot et al. 2005). In these examples, subtle regulation of gene function is achieved by regulating the ratio of spliced products.

Specific cases have been described, reporting splice forms with antagonist functions. One widely quoted example is the AS of Bcl-x, which encodes an antiapoptotic long form (Bcl-x(L)) and a proapoptotic short form (Bcl-x(S)).
Minn et al. (1996) have shown that 1 molecule of Bcl-x(S) per 4 molecules of Bcl-x(L) is necessary to overcome the survival mechanism of Bcl-x(L). Direct molecular interaction between isoforms, such as dominant-negative splice variants inhibiting functional variants by entering into non-functional heterodimers (McElvaine and Mayo 2006), could require quantitative regulation as well. Disequilibrium in the ratio between alternative splice variants can affect cell physiology, in some cases leading to disease (Nissim-Rafinia and Kerem 2002). Finally, a role for AS in regulating gene expression has been established at the mRNA level, associated to nonsense-mediated decay (NMD) (Lewis et al. 2003), and at the protein level, generating truncated proteins (Bingham et al. 1988). In this context, quantitative regulation can determine what proportion of transcripts leads to functional protein product. Up to 35% of all alternative isoforms are suggested to be targeted by NMD in human (Lewis et al. 2003).

The overall importance of the quantitative aspects of AS and its underlying regulation are still unclear, and very little is known about the evolutionary conservation of AS outside mammalian model systems. To assess the qualitative and quantitative evolutionary conservation of AS, we have determined the relative expression levels of AS variants from 21 selected pairs of orthologous AS events in Caenorhabditis elegans and Caenorhabditis briggsae through 6 developmental stages. These 2 nematode species split 80–110 MYA (Stein et al. 2003), about the same time as the human–mouse split. The nematode lineage shows a higher rate of evolutionary change than mammals with high levels of genome rearrangements, substitutions, gene duplications, and intron gain and loss (Stein et al. 2003). From this, one may assume that AS would be less conserved in nematodes than mammals, but our results suggest the opposite that AS seems to be highly conserved between C. elegans and C. briggsae.

Materials and Methods

Worm Growing

Worm strains were kept as previously described by Brenner (1974). The C. elegans strain used in this study was Bristol N2 (wild type). The C. briggsae strain was AF16 (wild type). All nematode strains used in this work were supplied by the Caenorhabditis Genetics Center, which is funded by the National Institutes of Health National Center for Research Resources. Eggs were obtained by treating adults (C. briggsae). All samples were checked for stage-specific anatomical characteristics as described by Wood (1988) and stored in RNAlater (Ambion, Inc., Austin, TX) for later RNA extraction. For each stage, 3 independent replicates were produced. All worms were cultured at 20 °C.

### Table 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene function summary</th>
<th>Type of AS event</th>
</tr>
</thead>
<tbody>
<tr>
<td>mig-15</td>
<td>Serine/threonine protein kinase</td>
<td>3’ trunc in ret int ret ex ret</td>
</tr>
<tr>
<td>eff-1</td>
<td>Transmembrane glycoprotein</td>
<td>3’ trunc</td>
</tr>
<tr>
<td>als-1</td>
<td>Predicted signal transduction protein</td>
<td>ex sk (5’)</td>
</tr>
<tr>
<td>deb-1</td>
<td>Vinculin ortholog</td>
<td>ex sk</td>
</tr>
<tr>
<td>cle-1</td>
<td>Collagen XV</td>
<td>ex sk (5’)</td>
</tr>
<tr>
<td>tie-1</td>
<td>Inositol triphosphate receptor</td>
<td>ex sk</td>
</tr>
<tr>
<td>gei-4</td>
<td>Intermediate filament interacting protein</td>
<td>ex sk</td>
</tr>
<tr>
<td>gfp-2</td>
<td>Intermediate filament</td>
<td>ex sk</td>
</tr>
<tr>
<td>unc-53</td>
<td>Axon guidance</td>
<td>ex sk (3’)</td>
</tr>
<tr>
<td>unc-52</td>
<td>Extracellular matrix structural protein</td>
<td>ex sk</td>
</tr>
<tr>
<td>unt-3</td>
<td>Troponin</td>
<td>ex sk</td>
</tr>
<tr>
<td>lin-10</td>
<td>Receptor clustering/synaptic transmission</td>
<td>ex sk</td>
</tr>
<tr>
<td>hil-7</td>
<td>Histone/nucleosome assembly</td>
<td>5’ trunc</td>
</tr>
<tr>
<td>uvt-5</td>
<td>Transferase activity/embryonic and larval development</td>
<td>ex sk</td>
</tr>
<tr>
<td>cup-5</td>
<td>Ion channel/apoptosis</td>
<td>ex sk</td>
</tr>
<tr>
<td>rne-8</td>
<td>Chaperone binding/endocytosis</td>
<td>ex sk</td>
</tr>
<tr>
<td>swp-1</td>
<td>Splicing factor</td>
<td>ex sk</td>
</tr>
</tbody>
</table>

Note.—When more than one event of a certain type occurs in the same gene, the event in case is identified by being closest to the 3’ (3’) or 5’ (5’) of the gene. ex sk, exon skipping; int ret, intron retention; trunc, truncation.

Selection of AS Events

A total of 21 AS events from 17 different genes are included in this study. The events were randomly selected from previously described AS events in annotated genes in the Wormbase database (www.wormbase.org). The C. briggsae orthologs were determined by BlastP best reciprocal match. If this did not provide an ortholog, the C. elegans sequence was blasted against the C. briggsae genome using BlastN, providing an ortholog if one existed. All genes are 1:1 reciprocal best matches. All pairs are confirmed in Wormbase homology link (based on best reciprocal matches and synteny alignments).

As a measure of the degree of which the sample genes are representative for the 2 species, orthologs were aligned using ClustalW and Ks/Ks ratios were calculated by using the Yang–Nielsen maximum-likelihood method, implemented in the YN00 program of the PAML package (Yang 1997). The complete set of splice events including type, gene function, and size of the alternatively spliced sequence is listed in table 1.

RNA Extraction and cDNA Synthesis

Total RNA was extracted from worms with the RNeasy Mini kit (Qiagen, Hilden, Germany) according to the recommendations of the manufacturer and included on-column DNase treatment using the RNase Free DNase Set (Qiagen). Reverse transcription–polymerase chain reaction (RT–PCR) was performed using a 19 nt long polyT primer with a nonthymidine anchor base at the 3’ end and SuperScriptIII RT–PCR kit (Invitrogen, Carlsbad, CA) according to the manufacturers’ instructions. The RT–PCR product was purified using the QIAquick PCR Purification kit (Qiagen) as specified by the manufacturer.
Multitemplate-Fluorescent PCR

To amplify alternative splice forms, primers annealing to the exons flanking the AS events were designed. Because the primers span the entire AS event, all isoforms from the event will be amplified (supplementary fig. S1, Supplementary Material online). Forward primers were fluorescently tagged at the 5’ end with FAM for C. elegans–specific primers and JOE for C. briggsae primers. PCR was performed using 0.8 ng/μl of cDNA template and 0.8 μM of each primer and 0.05 U/μl AmpliTaq Gold with ionic standard conditions. Elongation time was set to 6 min to avoid conditions that might favor amplification of short splice forms, and 27 cycles of PCR amplification were performed. That the method can accurately reflect transcript abundance is shown in supplementary figure S2 (Supplementary Material online), where splice variant frequencies are measured, reamplified, and then remeasured to check for possible amplification bias. Primer sequences and annealing temperatures used in this study are listed in supplementary table S1 (Supplementary Material online) is given.

Quantification of Alternative Splice Variants

PCR product (0.5 μl) was mixed with 1.5 μl ROX 2500 size standard mix (1× ROX2500 [Applied Biosystems, Foster City, CA], 2× ABI loading buffer [Applied Biosystems], and 6.75× formamide) and run under denaturing conditions in a 5% Polyacrylamide gel on an ABI PRISM 377 DNA Sequencer (Applied Biosystems). Electropherograms obtained were then analyzed using GeneScan Analysis Software version 3.1 (Applied Biosystems) and Genotyper 2.5 (Applied Biosystems). Peaks corresponding to the different isoforms were identified and their relative ratio calculated from peak area as described in the ABI Prism GeneScan Analysis Software User’s Manual (supplementary fig. S1, Supplementary Material online).

Statistical Analysis

To test whether AS relative expression levels change between subsequent stages and whether they change through development, a single-factor analysis of variance (ANOVA) was performed. A two-factor ANOVA with replication was performed to compare expression patterns from the two species. The “identity P value” between curves was used to determine whether the 2 curves were likely to be overlapping. Also, interaction effects were analyzed. If no significant interaction was found between the curves, they were considered as having parallel courses through development.

The change in the relative frequency of the isoforms, Δf, was calculated from the formula \( \Delta f_{sp} = f_{sp(i+1)} - f_{sp(i)} \), where \( f_{sp(i)} \) is the frequency of the longest isoform of a given gene in one of the 5 first developmental stages in Caenorhabditis development and \( f_{sp(i+1)} \) the frequency of the same isoform in the following stage. To avoid some data points having a disproportionate weight in statistical analysis, \( \Delta f \) values higher than 40% were considered statistical outliers and therefore discarded from analysis using this variable. A total of 4 points, corresponding to the changes in cle-1 (5’ ss) in the transition from embryos to L1 and unc-52 in the transition from L1 to L2 for both species, were discarded.

If the \( \Delta f \) distribution did not fit a normal distribution, a nonparametric analysis (Kruskall–Wallis test) was used instead of ANOVA. Throughout, \( P \) values <0.05 were considered significant.

Results

Strong Qualitative Conservation of AS Events and Isoforms

We have investigated 21 pairs of orthologous splice events in C. elegans and C. briggsae. All were known from Wormbase (www.wormbase.org) to be alternative splice events in C. elegans, yielding multiple mRNA products. A Ka/Ks calculation for the sample genes indicates that the conservation level of the selected genes can be considered representative for the 2 species, with an average Ka/Ks value of 0.06 identical to the overall C. elegans/C. briggsae average (Stein et al. 2003). For 19 of the 21 (90.5%) AS events in C. elegans, the AS event was also found in C. briggsae ortholog. The 2 exceptions are the alternative intron inclusion of intron 6 in mig-15 and the skipping of exon 5 in cup-5.

In C. elegans, 3 of the splice events (unc-53, hil-7, and wrt-5) have an output of 3 distinct splice variants, whereas the remaining 18 events give rise to 2 isoforms. The total output of the 21 events is therefore 45 splice variants. Of these, 42 (93.3%) splice variants were found to be conserved in C. briggsae.

Previously undescribed splice variants were cloned, sequenced, and submitted to GenBank (accession numbers DQ482052–93). In all cases, the sequences confirm that the splice variants identified in C. briggsae correspond to exact orthologs in C. elegans, with conserved splice donor and acceptor sites.

Most AS Variants Are Temporally Regulated through Worm Life Cycle

For every splice variant, a developmental expression profile was produced showing its frequency in 6 developmental stages (embryos [E], the 4 larval stages [L1–L4], and gravid adults [A]) (fig. 1). For each stage, measurements of 3 independent worm samples were obtained for each of the 2 species and the frequencies of the splice variants were determined as described in Materials and Methods. For all splice events, the frequency of the longest splice variant (i.e., the inclusion level of the alternatively spliced sequence) is shown for both species in figure 1. Interestingly, the frequencies of most of the splice variants are regulated through development. A single-factor ANOVA of the developmental profiles shows that in ~78% of all investigated AS events, relative isoform frequencies vary in a developmentally dependent manner (supplementary table S2, Supplementary Material online).

AS Isoform Frequencies Vary More in the First Developmental Stages

To determine whether variation of AS isoform frequency in the events analyzed is more common between some stages than others, we tested each AS event individually
Changes in the relative fraction of splice variants were found to take place more often in the first and last transitions in worm development. In the transition from embryos to L1 larvae, the frequency of 58% of all splice variants change, whereas 44% change between L4 larvae and gravid adults, whereas statistically significant changes between the individual larval stages were observed in only 5–25% of the cases.

To more thoroughly address how much variation is seen in the individual developmental transitions, we calculated the frequency change in percent for each transition, Δf (see Materials and Methods). The absolute values of Δf for each species (|Δf|) were plotted in a box-and-whiskers graph, grouping them in the 5 developmental transitions, as it is shown in figure 2A. We find that the average change in percent is higher in the first and last transitions for both species (fig. 2B). Interestingly, the patterns obtained for the 2 species are near identical, suggesting that the variation through development may be similar in both species.

A nonparametric analysis of these data shows that the E–L1 transition in both species and the L4–A transition in *Caenorhabditis elegans* have significantly higher isoform frequency variation. The plot of the real values of Δf (fig. 2C) shows that whereas inclusion level tends to decrease in embryos–L1, this is reversed in the L4–adults transition. This could be due to an effect of the developing embryos in gravid adults. Unlaid eggs could therefore be responsible for the high number of significant changes observed from L4 larvae to adults.

**Expression Profiles Show Similar Developmental Variation of AS Frequencies in *C. elegans* and *C. briggsae***

Expression profiles show similar development variation in *C. elegans* and *C. briggsae*. A nonparametric analysis of these data shows that the E–L1 transition in both species and the L4–A transition in *C. elegans* have significantly higher isoform frequency variation. The plot of the real values of Δf (fig. 2C) shows that whereas inclusion level tends to decrease in embryos–L1, this is reversed in the L4–adults transition. This could be due to an effect of the developing embryos in gravid adults. Unlaid eggs could therefore be responsible for the high number of significant changes observed from L4 larvae to adults.

Expression profiles show similar development variation of AS frequencies in *C. elegans* and *C. briggsae*. An AS isoform can be defined as functional if it is expressed in a regulated manner (Sorek et al. 2004). Nevertheless, variation in AS output through development may have no functional importance. In this case, such fluctuations would be unlikely to be conserved in evolution.
To address this issue, we estimated for each conserved AS event whether the associated expression profile is conserved as well. Using a two-factor ANOVA, no significant differences in the expression profiles were found in 8 of the 19 conserved orthologous events (42%). Moreover, even in the cases where the frequencies differ between the 2 species, the curves are often parallel (fig. 1). In 89% of the events, no interaction between the curves can be found, indicating that even if there is a measurable shift in splice variant frequency, the curves are parallel and the frequencies vary in a similar manner through development.

To assess the extent of conservation of developmental variation, a regression analysis was performed plotting all Δf for individual developmental transitions from *C. elegans* versus the ones from *C. briggsae* (fig. 3). This analysis shows that changes in AS isoform expression during development is strongly and significantly correlated between the 2 species ($r = 0.879$, $P < 0.0001$).

Thus, it seems that the majority of the splice events that we have investigated show conserved regulation during development. This suggests that the AS events are specifically regulated and also that the precise regulation of the AS isoforms is of functional relevance. To demonstrate the regulation and functionality of the individual isoforms, further experimental verification is warranted.

**Discussion**

Despite their morphological (Nigon and Dougherty 1949) and ecological similarity (Kiontke and Sudhaus 2006), the time of divergence of *C. elegans* and *C. briggsae* has been estimated to 80–110 MYA (Stein et al. 2003). Evolutionary conservation of a feature for such a long period of time supports a functional importance. In consequence, because both major and minor forms are conserved for most genes included in this study, both forms are likely to have functional importance. This is in contrast to earlier reports comparing mouse and human (Kan et al. 2002; Modrek and Lee 2003; Thanaraj et al. 2003). Also, in worms, because their relative expression levels are highly conserved, not only the presence of minor forms as well as major forms seems to be necessary but also the regulated proportion and/or location in which these are expressed. Moreover, our results suggest that the major versus minor form status was largely established before the 2 species diverged, although convergent evolution cannot be excluded.
However, the same isoform can have a frequency of 50% in some stages and 50% in others, which suggests that this categorization is developmentally dependent and to some extent arbitrary. For most (78%) of our set of AS events, the relative expression levels vary through development. These changes in the output of an AS event could reflect either that different splice variants are expressed in different developing tissues (gonad-specific variants would for instance increase greatly in frequency in late development) or, alternatively, that the AS output changes through the development of a single tissue. Both scenarios reflect the fact that AS is a highly regulated process. Highly correlated developmental variation in *Caenorhabditis elegans* and *Caenorhabditis briggsae* strongly suggests that AS regulatory mechanisms are also conserved in these 2 species. This is in agreement with recent findings, where conserved intronic AS cis-regulatory elements were identified by comparative genomics in nematodes (Kabat et al. 2006). These include elements conserved in *unc-52* and *deb-1* genes, both included in our data set. Furthermore, both cis- and trans-acting regulatory elements (Yeo et al. 2004) and overall splice variant expression levels, as reflected in EST databases, have been found to be conserved between human and mouse (Baek and Green 2005; Kan et al. 2005).

In 89% of the conserved studied AS events, the proportion of splice variants changes in a similar manner through the development of *C. elegans* and *C. briggsae*. However, the same isoform can have a frequency of >50% in some stages and <50% in others, which suggests that this categorization is developmentally dependent and to some extent arbitrary.

For most (78%) of our set of AS events, the relative expression levels vary through development. These changes in the output of an AS event could reflect either that different splice variants are expressed in different developing tissues (gonad-specific variants would for instance increase greatly in frequency in late development) or, alternatively, that the AS output changes through the development of a single tissue. Both scenarios reflect the fact that AS is a highly regulated process. Highly correlated developmental variation in *C. elegans* and *C. briggsae* strongly suggests that AS regulatory mechanisms are also conserved in these 2 species. This is in agreement with recent findings, where conserved intronic AS cis-regulatory elements were identified by comparative genomics in nematodes (Kabat et al. 2006). These include elements conserved in *unc-52* and *deb-1* genes, both included in our data set. Furthermore, both cis- and trans-acting regulatory elements (Yeo et al. 2004) and overall splice variant expression levels, as reflected in EST databases, have been found to be conserved between human and mouse (Baek and Green 2005; Kan et al. 2005).

In 89% of the conserved studied AS events, the proportion of splice variants changes in a similar manner through the development of *C. elegans* and *C. briggsae*.
Interestingly, the gene that showed the largest evolutionary developmental variation in alternative transcript abundance was itself a splicing factor, swp-1. Such variation might be expected to influence the splicing of genes affected by swp-1, in which case such genes might show different patterns of AS conservation/divergence.

The difference found in conservation of frequency variation and of the actual frequencies (i.e., that the curves are parallel) could be due to various factors. First, considering that we analyze RNA isolated from whole animals, it may be a consequence of a tissue-specific expression of the isoforms. If an isoform is expressed in a single tissue, a different relative size of this tissue between the 2 species would displace the curves in parallel. Second, it is possible that AS isoforms are coexpressed in the same cell type. Then the observed differences would reflect dynamic regulation within that cell. Finally, because AS is known to be environmentally regulated (Tao et al. 2002; Colot et al. 2005), the shift may be explained by the 2 species adapting differently to the laboratory environment.

The natural ecological niche of these species is thought to be highly similar (Kiontke and Sudhaus 2006), but it is possible that the similarities in AS regulation observed in this study are caused by analogous adaptation of these species to laboratory culturing. It would be interesting to test the effect of different environmental challenges on AS regulation of our gene set in both species. Results from such experiments would either strengthen the conclusion reached here for strong conservation of AS patterns or shed new light on ways that AS contributes to species-specific acclimation or stress responses.

Our data show that AS output varies most in the transition from developing embryos to L1 larvae. One could therefore speculate that AS is regulated differently in embryos compared with larval and adult stages. The fact that environment, morphology, and physiological functions are very different in eggs compared with larvae and adults supports this possibility.

Here, we have shown experimentally a high level of conservation of alternative splice variants as well as their relative frequency. To our knowledge, the level of conservation shown in this study is the highest yet reported. Most studies on AS conservation to date have been computational studies done in human–mouse systems. Although not quite consistent (Lareau et al. 2004), these generally estimate a low overall conservation level of 10–30% (Resch et al. 2004; Sorek et al. 2004; Sugnet et al. 2004; Pan et al. 2005; Yeo et al. 2005).

The divergence date for C. elegans and C. briggsae of 85–110 MYA (Stein et al. 2003) is similar to or earlier than the one estimated for the mouse–human pair (65–75 MYA). It is therefore possible that the Caenorhabditis transcriptome has diverged less over the same or longer evolutionary time than is the case with mouse and human, assuming that our data are representative.

However, methodological differences between different studies can also be of importance. Lareau et al. (2004) have pointed out that annotated genes may be enriched for conserved AS events. Studying only these, as is the case in the present study, would then lead to an overestimation of conservation levels. On the other hand, it is also likely that computational studies based on EST evidence underestimate the amount of conserved AS. It has recently been shown that much of nonconserved AS occurs in testis and cell lines where a high level of cellular stress and high proliferation rates can be expected to influence the accuracy of the splicing machinery (Kan et al. 2005). Also, more than 50% of ESTs available in databases derive from cancer, cell lines, or tumor tissues (Baranova et al. 2001). These factors suggest that a high proportion of available ESTs is the result of aberrant rather than regulated splicing and that conservation studies based on EST evidence alone may therefore have underestimated the amount of conserved AS. Furthermore, EST data are derived from a mixture of normalized, nonnormalized, subtracted, and nonsubtracted libraries. Normalization and subtraction may alter the relative abundance of isoforms, biasing the representation of splice variants. In contrast, the present investigation is based solely on study of wild-type organisms.

In general, a high rate of evolutionary change has been reported for the C. elegans and C. briggsae genomes with high levels of genome rearrangements (Coghlan and Wolfe 2002; Stein et al. 2003) and substitutions (Stein et al. 2003) compared with mammal and Drosophila model systems. Only ~85% of the intron positions are conserved between the 2 species of Caenorhabditis (Kent and Zähler 2000), whereas this conservation rises to 99.9% between mouse and human (Roy et al. 2003). In contrast to this, assuming that our findings are representative, AS of conserved genes contributes little to interspecies diversity between these nematodes. Also, only 65% of C. elegans genes have 1:1 C. briggsae orthologs, whereas the corresponding number for mouse/human is 80% (Stein et al. 2003). Interestingly, an inverse correlation between the size of a gene’s family and its use of alternatively spliced isoforms has been found for human and mouse genes, suggesting that these 2 evolutionary mechanisms can be used interchangeably (Kopelman et al. 2005). These factors further indicate that AS may not be a major contributor to transcript variation between the 2 species, whereas related mechanisms, like gene duplication, may be more active in this respect.

Indeed, our analysis shows very little alternatively spliced exon creation/loss in the evolutionary history of C. elegans and C. briggsae, whereas in mammals, AS has been associated with exon creation/loss (Modrek and Lee 2003). In contrast, it has been shown that more than 50% of newborn gene duplicates in C. elegans have unique exons in one or 2 members of the duplicate pair (Katju and Lynch 2003). However, full characterization of the rates and patterns of exon creation and loss associated with AS in Caenorhabditis will have to await larger scale comparisons between species.

The abundant differences at the genomic level between the 2 nematodes came as a surprise because they are morphologically almost indistinguishable (Stein et al. 2003). Our work shows that variability of transcriptome composition generated by AS is low, not displaying the variation seen in the genomic parameters just described. This would suggest a possible link between the highly conserved AS and the near-identical morphology of the 2 species. On the other hand, Caenorhabditis species are highly reduced animals, and it could be argued that the AS events that have
not been lost along this reductive process are more likely to be functional and important for the organism and therefore conserved in evolution.

To our knowledge, our study is the first experimental study on AS conservation to systematically investigate the quantitative conservation of AS. Our results suggest that the AS process is tightly regulated through development and that the regulatory mechanisms are to a large extent conserved over about 100 Myr between C. elegans and C. briggsae. This indicates functionality of both major and minor forms and of the relative quantity in which they are expressed. Finally, our results suggest that the quantitative regulation of isoform expression levels is an intrinsic part of most AS events.

Supplementary Material

Supplementary tables S1–S3 and figures S1–S3 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was funded by a Carlsberg Foundation grant (21-00-0680). We thank Scott W. Roy for his valuable help in revising the manuscript, Dan Jeffares for his ideas on the experimental design, and Paul Gardner for his help in preparing the manuscript and data analyses. We also thank the reviewers for constructive comments.

Funding to pay the Open Access publication charges for this article was provided by a Carlsberg Foundation Grant (21-00-0680).

Literature Cited


Douglas Crawford, Associate Editor

Accepted December 18, 2006