Alternative Splicing of the Drosophila Dscam Pre-mRNA Is Both Temporally and Spatially Regulated

Alicia M. Celotto and Brenton R. Graveley
Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06030

Manuscript received May 31, 2001
Accepted for publication July 6, 2001

ABSTRACT
The Drosophila melanogaster Down syndrome cell adhesion molecule (Dscam) gene encodes an axon guidance receptor that can express 38,016 different mRNAs by virtue of alternative splicing. The Dscam gene contains 95 alternative exons that are organized into four clusters of 12, 48, 33, and 2 exons each. Although numerous Dscam mRNA isoforms can be synthesized, it remains to be determined whether different Dscam isoforms are synthesized at different times in development or in different tissues. We have investigated the alternative splicing of the Dscam exon 4 cluster, which contains 12 mutually exclusive alternative exons, and found that Dscam exon 4 alternative splicing is developmentally regulated. The most highly regulated exon, 4.2, is infrequently used in early embryos but is the predominant exon 4 variant used in adults. Moreover, the developmental regulation of exon 4.2 alternative splicing is conserved in D. yakuba. In addition, different adult tissues express distinct collections of Dscam mRNA isoforms. Given the role of Dscam in neural development, these results suggest that the regulation of alternative splicing plays an important role in determining the specificity of neuronal wiring. In addition, this work provides a framework to determine the mechanisms by which complex alternative splicing events are regulated.

Drosophila melanogaster has ~250,000 neurons that are connected in a similar manner in each individual animal. Neurogenesis involves the migration of axons, which often occurs over long distances and through complex environments, to their targets. It is thought that the pattern of neuronal wiring is determined by a molecular code that involves specific interactions between receptor molecules on a given neuron and extracellular ligands (Albright et al. 2000). The receptors that participate in this process can be grouped into at least two classes, axon guidance receptors and synaptic cell adhesion molecules (Tessier-Lavigne and Goodman 1996).

The D. melanogaster cell adhesion molecule (Dscam) gene (Schmucker et al. 2000), which encodes an axon guidance receptor with similarity to the human Down syndrome cell adhesion molecule (Yamakawa et al. 1998), was recently identified by virtue of the interaction of its protein product with Dock, an adaptor protein involved in a signal transduction pathway required for normal axon guidance (Garrity et al. 1996). The Drosophila Dscam protein possesses an extracellular domain containing 10 immunoglobulin repeats and four fibronectin type III domains, which are connected to a transmembrane domain and a cytoplasmic domain (Schmucker et al. 2000). Dscam mRNA is expressed in neurons in both the central and peripheral nervous systems (Schmucker et al. 2000). Dscam mutants, which are early larval lethal, have axon guidance defects in the embryonic ventral nerve cord and Bolwig’s nerve (Schmucker et al. 2000). Together, these results suggest that Dscam encodes an axon guidance receptor that plays an important role in specifying neuronal wiring.

An extraordinary feature of the Dscam gene is that it can potentially produce 38,016 different mRNAs by virtue of alternative splicing (Black 2000; Schmucker et al. 2000; Graveley 2001). The Dscam gene contains 115 exons, 20 of which are constitutive and 95 of which are alternatively spliced (Figure 1). The alternatively spliced exons are organized into four clusters. The exon 4, 6, 9, and 17 clusters contain 12, 48, 33, and 2 variable exons, respectively. The exons within each cluster are alternatively spliced in a mutually exclusive manner such that Dscam mRNAs contain only one variable exon from each cluster. Alternative splicing of exons 4, 6, and 9 alters the encoded protein sequence of three extracellular immunoglobulin repeats. It is possible that different Dscam isoforms may interact with distinct sets of axon guidance cues. Alternative splicing of the Dscam pre-mRNA may therefore be central to the mechanisms specifying neuronal wiring. Thus, understanding the mechanisms of Dscam alternative splicing may provide insight into the genetic basis of neurogenesis in addition to providing an important model for investigating the regulation of complex alternative splicing events.

To begin to understand Dscam pre-mRNA alternative splicing in greater detail, we determined whether the...
alternative splicing of this gene is regulated. For the purposes of this study we focused our attention on the exon 4 cluster, which contains 12 variable exons. We find that the alternative splicing of exon 4 is regulated throughout development and in a tissue-specific manner. We also cloned and sequenced the exon 4 cluster of the Dscam gene from a related Drosophila species, D. yakuba, and found that the alternative splicing of D. yakuba Dscam exon 4 cluster is developmentally regulated. Products are first detectable, through cycle 40, which is beyond developmental time points, washed with PBS, and stored at −80°C until needed. RNA isolation: Total RNA was isolated using the LiCl-urea method (APFRAY and ROUGEOON 1980). Flies (0.2–0.5 g) at various stages of development were homogenized in 3 ml of a 3 M LiCl, 6 M urea solution and the homogenate was stored at −20°C overnight. The homogenate was removed and centrifuged at 13,000 rpm for 5 min. The pellet was resuspended in 300 μl of 1× TE (10 mM Tris, pH 8.0, 1 mM EDTA)/0.1% SDS. The solution was extracted twice with phenol-chloroform and the RNA was precipitated by adding 7.5 μl of 4 M NaCl and 750 μl of 100% ethanol.

To examine the tissue-specific profile of Dscam exon 4 expression, tissues (legs, wings, antennae, and heads) dissected from five animals were ground with a pestle in a microcentrifuge tube containing 300 μl of a 3 M LiCl, 6 M urea solution and 50 μg of yeast RNA as a carrier. After storing the solution overnight at −20°C the tube was then centrifuged at 13,000 rpm for 5 min and the pellet was resuspended in 50 μl of 1× TE/0.1% SDS. The RNA solution was extracted twice with phenol-chloroform and precipitated by adding 1 μl of 4 M NaCl and 250 μl of 100% ethanol.

**Primer:**

- Dmexon3us: 5′-TGCGACCCAAAAAGGACCG-3′
- Dmexon5ds: 5′-ACCGATGTCGACTCTCCAG-3′
- yakuba3: 5′-AGGACCCGGTTTCTTTCAGG-3′
- yakuba5: 5′-CAGAGGGCAAATCCAGAGAG-3′
- DSCAMex5rev: 5′-CGGATGTCGAGCTCTTCAAGGAGCC-3′

**Reverse transcription-polymerase chain reaction:** Reverse transcription (RT) reactions contained 20 units RNase inhibitor (Promega, Madison, WI), 200 units of SuperScript II (GIBCO BRL, Gaithersburg, MD); 5 μg of total fly RNA or the entire preparation of tissue RNA, 500 ng of random hexamers, and buffers provided by the manufacturer. The reaction was carried out at 42°C for 1 hr. Three microliters of the RT reaction was used as a template for polymerase chain reaction (PCR) using primers that anneal to exons 3 and 5 (D. melanogaster, Dmexon3us and Dmexon5ds; D. yakuba, yakuba3 and yakuba5). In each case, the exon 3 primer was end labeled with [γ-32P]ATP and T4 polynucleotide kinase. PCR was carried out for 45 sec at 94°C, 45 sec at 55°C, and 1 min at 72°C for 35 cycles using Taq DNA polymerase (Promega). Initial experiments were performed to determine whether the observed frequency that each exon is utilized fluctuates with the number of PCR cycles. A sampling of the products every 2 cycles revealed that although the total amount of product increased with each cycle, no significant differences were observed in the extent of amplification of each of the 12 Dscam RT-PCR products with respect to one another is relatively unaffected by cycle number. We believe that this is due to the fact that all of the products are the same size, amplified by the same primers, and therefore equally subject to saturation in the reaction.

**Single-strand conformation polymorphism gel electrophoresis:** To detect PCR products containing each of the exon 4 variants, we used single-strand conformation polymorphism (SSCP)/multidetection enhancement (MDE) gels. The SSCP gels (20 cm × 45 cm × 0.4 mm) contained 0.6× TBE and 25% MDE gel solution (BioWhittaker Molecular Applications, Walkersville, MD). Ten microliters of stop solution (95% formamide, 1 M NaOH, 0.25% bromophenol blue, 0.25% xylene cyanol) was added to 2 μl of each PCR reaction. This solution was heated to 95°C for 2 min, placed on ice for 5 min, and 2 μl of each sample was loaded onto the gel. Each SSCP gel was run for 24 hr at a maximum of 8 W and a constant temperature of 25°C. The gel was dried and exposed to film with an intensifying screen at −80°C overnight. Quantitation was performed using a Molecular Dynamics (Sunnyvale, CA) Storm phosphorimager and ImageQuant software.

**Cloning of Dscam cDNA fragments:** RT-PCR products obtained using the Dmexon3us and Dmexon5ds (D. melanogaster) or yakuba3 and yakuba5 (D. yakuba) primers were cloned into the pCRII-TOPO-TA vector (Invitrogen, San Diego) as described by the manufacturer. Clones were picked by random and sequenced using the DSCAMex5rev primer (D. melanogaster) or yakuba5 (D. yakuba) to identify one clone containing each variant of exon 4. These clones were used as templates for PCR using the primers Dmexon3us and Dmexon5ds (D. melanogaster) or yakuba3 and yakuba5 (D. yakuba) to generate a set of SSCP migration standards.

**Cloning the exon 4 region of the D. yakuba Dscam gene:** D. yakuba genomic DNA was prepared from 50 adult flies using the QIAamp DNA mini kit (QIAGEN, Valencia, CA). The D. yakuba Dscam exon 4 region was amplified from 300 ng D. yakuba genomic DNA, using the Dmexon3us and Dmexon5ds primers. The PCR reaction was carried out for 1 min at 94°C, 1 min at 55°C, and 15 min at 72°C for 35 cycles using 2.5 units of LA Taq (TaKaRa, Berkeley, CA). The PCR product was cloned into the pCRII-TOPO-TA vector (Invitrogen) as described by the manufacturer. The sequence of the exon 4 region of the D. yakuba Dscam gene was determined from two independent clones. Automated sequencing was performed at the University of Connecticut Health Center Molecular Core Facility. The sequence of the D. yakuba Dscam exon 4 region has been deposited in GenBank (accession no. AF389590).

**RESULTS**

**Analysis of Dscam exon 4 alternative splicing:** The Drosophila Dscam gene contains 95 alternative exons arranged in four clusters (Figure 1). The exon 4, 6, 9, and 17 clusters contain 12, 48, 33, and 2 variable exons,
respectively. We are interested in understanding the alternative splicing of the *Dscam* transcripts in greater detail and in particular whether the alternative splicing is regulated. For the purposes of these studies, we focused on the alternative splicing within the exon 4 cluster.

All of the exon 4 variants are very similar in size, ranging from 159 to 171 nucleotides (nt; 4.1 11 = 159 nt; 4.1, 4.2, 4.3, 4.5, 4.6, 4.7 = 162 nt; 4.9 = 168 nt; and 4.4, 4.8, 4.10, 4.12 = 171 nt). Likewise, the 12 possible RT-PCR products obtained using primers in the constant exons 3 and 5 differ from one another by only 12 nt. As a result, traditional methods such as agarose gel electrophoresis cannot be used to analyze the alternative splicing of exon 4. By separating the RT-PCR products on a SSCP gel, which separates molecules based on conformational differences (Orita *et al.* 1989), we were able to distinguish the majority of the 12 RT-PCR products from one another (Figure 2). The identity of each band was assigned in two ways. First, standards were generated using PCR products generated from 12 cDNA clones containing each exon 4 variant spliced to exons 3 and 5 (lanes 2–13). Second, each band from a reaction similar to that shown in lane 1 was excised from the SSCP gel, cloned, and sequenced.

Although the majority of the RT-PCR products migrate at a distinct position in the gel, the RT-PCR products obtained from some mRNAs comigrate. Specifically, RT-PCR products containing exons 4.3 and 4.12 migrate as a single band as do the RT-PCR products synthesized from mRNAs containing exons 4.5, 4.7, and 4.9. Using this method, we can determine the relative frequency with which the majority of the exon 4 variants are utilized within each RNA sample. The percentage inclusion of each exon is calculated by dividing the total pixels in each individual band by the sum of the pixels in all bands in a single lane.

**Developmentally regulated alternative splicing of**

**Dscam exon 4 in *D. melanogaster***: To determine if *Dscam* alternative splicing is regulated, we measured the frequency at which each *Dscam* exon 4 variant is utilized throughout development. Total RNA harvested from flies at various stages of development was used as a tem-
plate for RT-PCR reactions with primers in exons 3 and 5 and the RT-PCR products were resolved on SSCP gels (Figure 3). The frequency at which most of the exon 4 variants are utilized does not change significantly throughout development. However, the splicing of two exons, 4.2 and 4.8, appears to be highly regulated (Table 1).

Exon 4.2 displays the most striking developmental changes. Only \( \approx 1\% \) of the \textit{Dscam} transcripts in 0- to 12-hr embryos contain exon 4.2. However, in first instar larvae (L1), exon 4.2-containing transcripts make up \( \approx 20\% \) of the total \textit{Dscam} mRNAs. An analysis of RNA isolated hourly from embryos raised at 22°C revealed that \textit{Dscam} transcripts containing exon 4.2 first appear at hour 12, which corresponds to embryonic stage 15 (data not shown). The relative abundance of \textit{Dscam} transcripts containing exon 4.2 remains high throughout the remainder of development and, in adults, \( \approx 44\% \) of the total \textit{Dscam} mRNAs contain exon 4.2. This represents a 40- to 50-fold increase in exon 4.2 utilization between embryos and adults.

The expression of \textit{Dscam} mRNAs containing exon 4.8 is the opposite of the expression pattern of exon 4.2-containing transcripts. Approximately 20% of all \textit{Dscam} mRNAs in embryos contain exon 4.8. The abundance of exon 4.8-containing transcripts decreases throughout the remainder of development and in adults only \( \approx 1\% \) of the total \textit{Dscam} mRNAs contain exon 4.8. We conclude that alternative splicing of some of the \textit{Dscam} exon 4 variants is dramatically regulated throughout development.

**Tissue-specific alternative splicing of \textit{Dscam} exon 4 in \textit{D. melanogaster}:** The diversity of \textit{Dscam} proteins generated by alternative splicing is thought to play an important role in determining the specificity of neuronal wiring (Schmucker et al. 2000). One prediction of this model is that neurons in different tissues would express different \textit{Dscam} isoforms to direct their axons to specific addresses. To begin testing this model, we examined the relative abundance of each exon 4-containing \textit{Dscam} mRNA isoform in different adult tissues.

RNA was harvested from antennae, heads, wings, and legs dissected from adult flies. These RNA samples were subjected to RT-PCR and the products were separated on SSCP gels (Figure 4). These results show that the collection of \textit{Dscam} transcripts is significantly different in each body part examined. For example, whereas \( \approx 45\% \) of the total \textit{Dscam} transcripts in legs and \( \approx 42\% \) of the transcripts in wings contain exon 4.2, only \( \approx 16\% \) of \textit{Dscam} transcripts isolated from heads contain exon 4.2 (lanes 2–4). Similar differences are observed for many of the other exon 4 variants (see Table 2). We conclude that the alternative splicing of the \textit{Dscam} exon 4 variants is regulated in a tissue-specific manner.

**Cloning and analysis of the exon 4 region of the \textit{D. yakuba} \textit{Dscam} gene:** The most striking change we observed is the developmental regulation of exon 4.2, which is not utilized in early embryos. To determine
### TABLE 1
Developmental regulation of Dscam exon 4 alternative splicing in *D. melanogaster*

<table>
<thead>
<tr>
<th>Exon</th>
<th>Embryo 0–3 hr</th>
<th>Embryo 3–6 hr</th>
<th>Embryo 6–9 hr</th>
<th>Embryo 9–12 hr</th>
<th>L1</th>
<th>L2</th>
<th>L3 non</th>
<th>L3 wand</th>
<th>Pupae</th>
<th>Pharate</th>
<th>Adult</th>
<th>Virgins</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>11.6 ± 1.0</td>
<td>11.5 ± 1.6</td>
<td>11.2 ± 2.5</td>
<td>12.7 ± 3.7</td>
<td>10.0 ± 0.6</td>
<td>7.6 ± 3.2</td>
<td>12.8 ± 2.4</td>
<td>16.0 ± 1.8</td>
<td>16.1 ± 1.3</td>
<td>9.2 ± 1.3</td>
<td>12.9 ± 5.4</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>1.6 ± 1.4</td>
<td>0.7 ± 0.5</td>
<td>0.9 ± 0.7</td>
<td>0.5 ± 0.4</td>
<td>22.8 ± 1.4</td>
<td>18.7 ± 7.9</td>
<td>18.8 ± 0.8</td>
<td>15.0 ± 1.6</td>
<td>14.3 ± 1.2</td>
<td>31.7 ± 3.7</td>
<td>44.2 ± 6.6</td>
<td></td>
</tr>
<tr>
<td>4.3/12</td>
<td>16.6 ± 5.1</td>
<td>20.2 ± 3.5</td>
<td>19.8 ± 2.1</td>
<td>16.2 ± 3.8</td>
<td>12.5 ± 1.9</td>
<td>14.6 ± 1.9</td>
<td>12.6 ± 0.7</td>
<td>11.0 ± 0.9</td>
<td>11.2 ± 0.9</td>
<td>12.7 ± 1.4</td>
<td>11.1 ± 3.3</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>6.4 ± 0.4</td>
<td>7.6 ± 1.2</td>
<td>7.2 ± 0.9</td>
<td>6.3 ± 1.1</td>
<td>5.5 ± 1.5</td>
<td>7.7 ± 0.7</td>
<td>10.5 ± 2.0</td>
<td>11.0 ± 1.4</td>
<td>6.8 ± 1.0</td>
<td>4.4 ± 1.0</td>
<td>5.8 ± 2.4</td>
<td></td>
</tr>
<tr>
<td>4.5/7/9</td>
<td>8.3 ± 0.9</td>
<td>7.1 ± 0.7</td>
<td>8.4 ± 0.9</td>
<td>10.7 ± 1.4</td>
<td>13.0 ± 1.8</td>
<td>14.3 ± 7.6</td>
<td>11.4 ± 2.3</td>
<td>12.3 ± 2.4</td>
<td>8.1 ± 0.6</td>
<td>9.3 ± 1.2</td>
<td>10.9 ± 5.1</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>16.3 ± 2.0</td>
<td>13.6 ± 1.1</td>
<td>14.4 ± 3.3</td>
<td>14.2 ± 1.4</td>
<td>9.5 ± 0.3</td>
<td>6.9 ± 0.9</td>
<td>11.0 ± 2.0</td>
<td>11.1 ± 2.5</td>
<td>8.6 ± 1.3</td>
<td>7.1 ± 0.9</td>
<td>6.9 ± 1.8</td>
<td></td>
</tr>
<tr>
<td>4.8</td>
<td>21.8 ± 4.6</td>
<td>19.7 ± 5.5</td>
<td>18.3 ± 5.8</td>
<td>21.8 ± 4.0</td>
<td>12.2 ± 1.3</td>
<td>12.2 ± 5.3</td>
<td>12.7 ± 1.7</td>
<td>11.5 ± 1.9</td>
<td>17.1 ± 2.0</td>
<td>10.7 ± 4.4</td>
<td>1.3 ± 2.4</td>
<td></td>
</tr>
<tr>
<td>4.10</td>
<td>6.7 ± 1.5</td>
<td>7.4 ± 1.2</td>
<td>6.8 ± 0.6</td>
<td>8.7 ± 0.6</td>
<td>7.7 ± 0.5</td>
<td>6.9 ± 1.2</td>
<td>5.0 ± 1.1</td>
<td>6.3 ± 1.1</td>
<td>9.2 ± 1.7</td>
<td>8.3 ± 1.8</td>
<td>2.5 ± 1.5</td>
<td></td>
</tr>
<tr>
<td>4.11</td>
<td>10.8 ± 2.8</td>
<td>12.2 ± 1.0</td>
<td>13.3 ± 2.8</td>
<td>9.1 ± 2.4</td>
<td>7.0 ± 0.9</td>
<td>11.1 ± 0.5</td>
<td>5.3 ± 0.9</td>
<td>5.8 ± 1.4</td>
<td>8.5 ± 1.5</td>
<td>6.6 ± 0.6</td>
<td>4.4 ± 0.9</td>
<td></td>
</tr>
</tbody>
</table>

All data represent the percentage inclusion of each exon calculated as described in the Figure 3 legend. The standard deviation was calculated from at least three independent experiments. L3 non, L3 nonwandering; L3 wand, L3 wandering.
A comparison of the splice site sequences flanking each of the exon 4 variants revealed that the 5′ splice sites are more conserved between the two species than the 3′ splice sites (Table 3). For example, between the two species there are only 2 nucleotide changes among the 12 exon 4 5′ splice sites, whereas there are 25 nucleotide changes among the 12 exon 4 3′ splice sites.

The regulation of Dscam alternative splicing is conserved between D. melanogaster and D. yakuba: We next tested whether the developmental pattern of exon 4.2 alternative splicing observed in D. melanogaster is conserved in D. yakuba. The alternative splicing pattern of exon 4 throughout D. yakuba development was determined by performing RT-PCR, using D. yakuba-specific primers in exons 3 and 5 on total RNA harvested from embryos, larvae, and adults. The reactions were resolved on an SSCP gel to visualize the various exon 4-containing PCR products (Figure 6A). The migration pattern of the D. yakuba exon 4-containing PCR products differs from that of D. melanogaster. We therefore determined the identity of each band with standards derived from cloned D. yakuba Dscam exon 4-containing cDNAs (data not shown). In addition, each band was excised from the gel, PCR amplified, cloned, and sequenced. As with D. melanogaster, we find that D. yakuba Dscam transcripts containing exon 4.2 are not expressed in embryos but are expressed in both larvae and adults (Figure 6, A and B, Table 4). We also find that the relative abundance of the exon 4.8-containing transcripts decreases throughout D. yakuba development as in D. melanogaster (Figure 6C and Table 4). However, in both cases, the magnitude of the changes is lower in D. yakuba than in D. melanogaster (Figure 6, B and C). We conclude that a similar developmental pattern of Dscam exon 4-regulated alternative splicing occurs in both D. melanogaster and D. yakuba.

DISCUSSION

The Drosophila Dscam gene can potentially generate more alternatively spliced mRNAs than any other gene known in nature (Black 2000; Graveley 2001). Here, we examined whether the alternative splicing of this gene is regulated. In particular, we have focused on the exon 4 region of the Dscam gene, where the splicing machinery must select 1 of 12 possible alternative exons. These experiments demonstrated that the alternative splicing of this region is controlled throughout development and in a tissue-specific manner. In addition, we found that at least some of this regulated alternative splicing is evolutionarily conserved. This work lays the foundation to begin examining the biochemical mechanisms involved in controlling the extraordinarily complex alternative splicing of this gene, which plays an important role in determining the specificity of neuronal wiring.

Our most striking observation is the developmental regulation of exon 4.2 alternative splicing. This exon is

Table 2

<table>
<thead>
<tr>
<th>Exon</th>
<th>Antenna</th>
<th>Head</th>
<th>Wings</th>
<th>Legs</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>4.7 ± 0.4</td>
<td>12.3 ± 3.7</td>
<td>13.6 ± 4.3</td>
<td>11.5 ± 4.3</td>
</tr>
<tr>
<td>4.2</td>
<td>26.9 ± 6.8</td>
<td>15.8 ± 2.5</td>
<td>42.4 ± 11.5</td>
<td>44.8 ± 14.6</td>
</tr>
<tr>
<td>4.3/12</td>
<td>17.4 ± 11.1</td>
<td>22.8 ± 5.6</td>
<td>11.6 ± 5.9</td>
<td>9.2 ± 2.5</td>
</tr>
<tr>
<td>4.4</td>
<td>6.4 ± 2.8</td>
<td>3.6 ± 0.1</td>
<td>3.2 ± 1.6</td>
<td>2.7 ± 0.2</td>
</tr>
<tr>
<td>4.5/7/9</td>
<td>14.1 ± 5.8</td>
<td>12.4 ± 1.0</td>
<td>10.8 ± 10.0</td>
<td>7.6 ± 2.4</td>
</tr>
<tr>
<td>4.6</td>
<td>7.7 ± 3.8</td>
<td>10.3 ± 1.0</td>
<td>5.7 ± 1.0</td>
<td>8.7 ± 2.8</td>
</tr>
<tr>
<td>4.8</td>
<td>10.1 ± 4.7</td>
<td>9.9 ± 2.1</td>
<td>6.9 ± 1.9</td>
<td>8.4 ± 1.1</td>
</tr>
<tr>
<td>4.10</td>
<td>3.2 ± 1.6</td>
<td>5.9 ± 1.5</td>
<td>2.0 ± 0.8</td>
<td>3.4 ± 1.8</td>
</tr>
<tr>
<td>4.11</td>
<td>9.5 ± 3.7</td>
<td>6.9 ± 1.3</td>
<td>3.8 ± 0.4</td>
<td>3.7 ± 0.3</td>
</tr>
</tbody>
</table>

All data represent the percentage inclusion of each exon calculated as described in the Figure 3 legend. The standard deviation was calculated from three independent experiments.

Figure 5.—Comparison of the exon 4 region of the Dscam gene from D. melanogaster and D. yakuba. The exon 4 region of the Dscam gene was cloned from D. yakuba by PCR using primers in exons 3 and 5. The entire nucleotide sequence was determined and aligned to the sequence of the D. melanogaster Dscam gene using the program Advanced PipMaker (Schwartz et al. 2000). The data are shown as a percentage identity plot. The locations of the exons are shown. The height of the horizontal bars within the graph indicates the degree of homology between the two sequences.
The extraordinary diversity of Dscam isoforms is likely to be critical to its function as an axon guidance receptor. Presumably, individual neurons will be found to express different Dscam isoforms, the collection of which will determine the path that a neuron will take in finding its synaptic target. Consistent with this idea, we found that the alternative splicing of Dscam exon 4 is regulated in a tissue-specific manner. Although our experiments demonstrate that there are dramatic differences in the collection of Dscam isoforms expressed in the tissues we examined, we imagine that even greater differences will be observed when the Dscam expression profiles are examined at a more detailed level. This is due to the fact that the tissues we examined undoubtedly contain thousands to tens of thousands of neurons. It will be critical to determine how many, and which, Dscam mRNA isoforms are expressed in individual neurons.

Our comparison of Dscam exon 4 alternative splicing in D. melanogaster and D. yakuba demonstrates that both exon 4.2 and exon 4.8 are regulated in a similar manner in the two species. However, we also note that some exons are utilized at different frequencies in the two species. For example, in D. melanogaster exons 4.5, 4.7, and 4.9 (which migrate as a single band in the SSCP gels) together make up between 8 and 14.3% of the total Dscam transcripts throughout development (Table 1). In contrast, in D. yakuba, 12–23% of the total Dscam transcripts contain exon 4.5 and 13–16% contain exon 4.7 (Table 4). Thus, while there are similarities in how the alternative splicing of some of the exon 4 variants is regulated in a tissue-specific manner, there are also significant differences in the relative abundance of several Dscam mRNA isoforms. It is possible that such differences could underlie behavioral variations between the
two species. In addition, these differences could have arisen if the morphological differences of these two species require axons to travel along alternate routes to find a similar target.

Regulatory mechanisms of Dscam alternative splicing: Our results demonstrate that the alternative splicing of Dscam exon 4 is regulated throughout development and in a tissue-specific manner. What types of regulatory mechanisms may act in controlling the splicing of these exons? We can envision at least two models. The first is a positive regulatory model. In this case, each exon 4 variant would be skipped by default. The binding of a regulatory factor or complex at or near a specific exon would then activate the splicing of that exon. The second model involves negative regulation. In this case, each exon 4 variant would be included by default and the selection of a single exon would require the active repression of the other 11 exon 4 variants. Future work will be aimed at defining the biochemical mechanisms responsible for regulating exon 4 alternative splicing.

Comparative sequence analysis can be used to gain insight into important aspects of gene structure and may highlight potential cis-acting regulatory regions (Kent and Zahler 2000). We cloned and sequenced the exon 4 region of the Dscam gene from D. yakuba, which is estimated to have diverged from D. melanogaster ~7–15 million years ago (Powell 1997; Li et al. 1999; Robin et al. 2000). The sequence of the D. yakuba exon 4 region is quite similar to that of D. melanogaster throughout its length (Figure 5). The genes from both species contain 12 exon 4 variants and the spatial organization of the exons and introns is remarkably conserved (Table 3). The nucleotide and protein sequences of the alternative exons are highly conserved (>90% identity), while the intron sequences flanking the alternative exons are less similar (~80% identity). Unfortunately, this

### TABLE 4

<table>
<thead>
<tr>
<th>Exon</th>
<th>Embryo</th>
<th>Larvae</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>7.4 ± 0.7</td>
<td>7.5 ± 0.4</td>
<td>3.3 ± 0.2</td>
</tr>
<tr>
<td>4.4</td>
<td>8.7 ± 0.05</td>
<td>6.2 ± 0.8</td>
<td>6.8 ± 0.2</td>
</tr>
<tr>
<td>4.5</td>
<td>23.0 ± 2.4</td>
<td>21.5 ± 0.7</td>
<td>21.2 ± 0.9</td>
</tr>
<tr>
<td>4.6/10</td>
<td>9.8 ± 0.4</td>
<td>6.4 ± 0.7</td>
<td>4.8 ± 0.7</td>
</tr>
<tr>
<td>4.7</td>
<td>13.1 ± 0.4</td>
<td>13.5 ± 1.9</td>
<td>16.7 ± 0.5</td>
</tr>
<tr>
<td>4.8</td>
<td>14.7 ± 1.2</td>
<td>11.1 ± 0.01</td>
<td>9.8 ± 0.4</td>
</tr>
<tr>
<td>4.9/12</td>
<td>10.6 ± 1.0</td>
<td>10.9 ± 1.1</td>
<td>12.1 ± 2.6</td>
</tr>
<tr>
<td>4.11</td>
<td>12.6 ± 0.08</td>
<td>10.7 ± 1.0</td>
<td>10.9 ± 1.0</td>
</tr>
</tbody>
</table>

All data represent the percentage inclusion of each exon calculated as described in the Figure 3 legend. The standard deviation was calculated from two independent experiments.
level of conservation is too high to clearly identify important cis-acting regulatory elements. Determining the sequence of the Dscam exon 4 region from more distantly related flies will be necessary to identify potential splicing regulatory elements.

To determine whether the splicing of the exon 4 variants is regulated by controlling 5’ or 3’ splice site utilization, we compared the sequences of the splice sites flanking each alternative exon. This analysis revealed that the exon 4 5’ splice sites are more conserved between D. melanogaster and D. yakuba than the exon 4 3’ splice sites (Table 3). While the Drosophila 5’ splice site consensus sequence is rather constrained, the 3’ splice site consensus sequence is larger and more relaxed (Mount et al., 1992). Thus, the nucleotide differences we observe between the two species may simply reflect neutral evolutionary change. Alternatively, these changes may indicate that the regulation of exon 4 alternative splicing may be achieved through 5’ splice site selection instead of 3’ splice site selection. Regardless, the analysis of sequences from more distantly related species, coupled with biochemical experiments, will be required to elucidate the mechanisms controlling the alternative splicing of these exons.

A method for analyzing complex alternative splicing events: The unusual exon-intron structure and alternative splicing patterns of the Dscam gene required the development of a new method for analyzing alternative splicing. Given that the 12 exon 4 variants are nearly identical in size, methods such as agarose gel electrophoresis or denaturing polyacrylamide gels were not suitable for separating the RT-PCR products derived from the Dscam mRNAs. However, we found that SSCP gels were well suited for this analysis. These gels separate DNA molecules on the basis of their conformation, which is determined by their nucleotide sequence rather than their size (Orita et al., 1989). Although the Dscam gene appears to have the most complicated organization, there are numerous examples of eukaryotic genes that contain mutually exclusive alternatively spliced exons of the same size. The vast majority of the known examples contain only two exons (Stamm et al., 2000), but at least one, the Drosophila myosin heavy chain (Standiford et al., 1997, 2001), contains multiple tandemly arranged exons. In fact, like Dscam, the Drosophila myosin heavy chain gene contains three clusters, each containing multiple alternatively spliced exons. In this case, the exon 7 cluster contains 4 variants, the exon 9 cluster contains 3 variants, and the exon 11 cluster contains 5 variants (Standiford et al., 1997, 2001). The relatively simple method we developed could be effectively used to analyze the alternative splicing of transcripts from any gene containing mutually exclusive exons.

The complexity of alternative splicing: One of the most unexpected discoveries of the genomic era is that gene number does not correlate with the complexity of an organism (Black 2000; Graveley 2001). One solution to this paradox is that processes such as alternative splicing function to generate a tremendously diverse proteome from a relatively small number of genes. It is estimated that at least 55% of all human genes (Kan et al., 2001), and perhaps as many Drosophila genes, encode alternatively spliced transcripts and some of these are alternatively spliced to generate an amazingly large number of distinct mRNA isoforms (Black 2000; Graveley 2001). However, the amount of protein diversity generated by alternative splicing of the D. melanogaster Dscam pre-mRNAs is unparalleled (Schmucker et al., 2000). Understanding the mechanisms by which Dscam alternative splicing is regulated will provide insight into not only the basic mechanisms of regulated alternative splicing but also into how the specificity of neuronal wiring is genetically encoded.

We thank D. Black, J. Clemons, A. Das, E. Fleming, K. Hertel, K. Lynch, W. Mattox, M. Palladino, R. Reenan, L. Zipursky, and members of our laboratory for discussions and/or comments on the manuscript. This work was supported by startup funds and a Research Initiation and Support Enhancement (RISE) grant (to B.R.G.) from the University of Connecticut Health Center.

---

**LITERATURE CITED**


Communicating editor: A. J. Lopez