The *Caenorhabditis elegans* schnurri homolog *sma-9* mediates stage- and cell type-specific responses to DBL-1 BMP-related signaling

Jun Liang¹, Robyn Lints²,³, Marisa L. Foehr⁴, Rafal Tokarz¹, Ling Yu¹, Scott W. Emmons², Jun Liu⁴ and Cathy Savage-Dunn¹,*

¹Department of Biology, Queens College, and PhD Program in Biochemistry, The Graduate School and University Center, The City University of New York, Flushing, NY 11367, USA
²Department of Molecular Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA
³Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA
⁴Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

*Author for correspondence (e-mail: csavage@qc1.qc.edu)

Accepted 11 September 2003

**Summary**

In *Caenorhabditis elegans*, the DBL-1 pathway, a BMP/TGFβ-related signaling cascade, regulates body size and male tail development. We have cloned a new gene, *sma-9*, that encodes the *C. elegans* homolog of Schnurri, a large zinc finger transcription factor that regulates dpp target genes in *Drosophila*. Genetic interactions, the *sma-9* loss-of-function phenotype, and the expression pattern suggest that *sma-9* acts as a downstream component and is required in the DBL-1 signaling pathway, and thus provide the first evidence of a conserved role for Schnurri proteins in BMP signaling. Analysis of *sma-9* mutant phenotypes demonstrates that SMA-9 activity is temporally and spatially restricted relative to known DBL-1 pathway components. In contrast with *Drosophila schnurri*, the presence of multiple alternatively spliced *sma-9* transcripts suggests protein isoforms with potentially different cell sublocalization and molecular functions. We propose that SMA-9 isoforms function as transcriptional cofactors that confer specific responses to DBL-1 pathway activation.

Key words: BMP, Schnurri, Transcription factor, Body size, Pattern formation, Alternative splicing

**Introduction**

The bone morphogenetic proteins (BMPs) are members of the transforming growth factor β (TGFβ) superfamily of secreted peptide growth factors that play crucial roles in development and cell differentiation in both invertebrate and vertebrate model systems (Nakayama et al., 2000; Patterson and Padgett, 2000; Hogan, 1996). Signal transduction for this family of growth factors is mediated by two transmembrane Ser/Thr kinase receptors and two or three intracellular Smads (Massague, 1998). Smads function by forming complexes that shuttle into the nucleus and activate transcription of downstream target genes. How this simple canonical signaling cassette elicits specific responses is not well understood. In vitro, the Smad complex binds a low complexity DNA sequence with low affinity, which suggests that, in vivo, efficient promoter binding and target gene regulation may require interaction with transcriptional cofactors (Shi et al., 1998; Massague, 1998). However, to date, few Smad-interacting transcriptional cofactors have been identified: *Drosophila* Schnurri (Shn); and vertebrate FAST1, FAST2, OAZ and Mix (Arora et al., 1995; Grieder et al., 1995; Staehling-Hampton et al., 1995; Chen et al., 1996; Hata et al., 2000; Randall et al., 2002).

*Drosophila* Shn is required for the Dpp BMP-related pathway. Shn loss of function causes embryonic lethality and ventralization, as does loss of other pathway components. In vitro, Shn was demonstrated to interact with the Smad protein Mad, to bind specific DNA sequences, and to recognize a Dpp-responsive promoter element of the *Ubx* gene (Dai et al., 2000). However, there is still uncertainty about how Shn functions. One possibility is that it may regulate Dpp-mediated transcriptional activation of target genes directly (Torres-Vazquez et al., 2001), and another is that it may activate target genes indirectly by transcriptional repression of *brinker*, a novel transcriptional repressor (Campbell and Tomlinson, 1999; Jazwinska et al., 1999; Marty et al., 2000; Muller et al., 2003). There are Shn homologs present in vertebrate and *C. elegans* genomes. Vertebrate homologs include Shn1 (MBP-1/PRDII-BF1/αA-CRYBP1/GAAP-1/HIV-EP1), Shn2 (MBP-2/HIV-EP2) and Shn3 (HIV-EP3/KRC); however, no reports have shown their functions in BMP signaling (Fan and Maniatis, 1990; Nakamura et al., 1990; van’t Veer et al., 1992; Seeler et al., 1994; Gascoigne, 2001; Takagi et al., 2001; Lallemend et al., 2002; Oukka et al., 2002). The question arises as to whether these proteins play a conserved role in BMP signaling or not. We have used *C. elegans* as a model system to address this question.

In the nematode *C. elegans*, a BMP-related signaling pathway regulates body size and patterning of sex-specific tissues of the male posterior (Patterson and Padgett, 2000;...
Savage-Dunn, 2001). Studies of this pathway have previously been fruitful in identifying conserved signaling components (Savage et al., 1996). The pathway, which we will refer to as the DBL-1 pathway, is defined by six genes: ligand dbl-1 (Suzuki et al., 1999; Morita et al., 1999), type I receptor sma-6 (Krishna et al., 1999), type II receptor daf-4 (Estevez et al., 1993), and Smads sma-2, sma-3 and sma-4 (Savage et al., 1996). Mutations in any of these pathway components cause a small body size (Sma) phenotype in both hermaphrodites and males, and a male abnormal (Mab) phenotype due to transformations in male sensory ray identity and defective morphogenesis of the male copulatory spicules.

To identify additional components of the DBL-1 pathway and in particular those that confer specific pathway responses, we performed forward genetic screens for additional mutations affecting body size or male sensory ray patterning (Savage-Dunn et al., 2003; Lints and Emmons, 2002). Here we report the isolation of the C. elegans shn homolog sma-9, and provide the first evidence of a conserved role for Shn proteins in BMP-related signaling. Loss-of-function (lf) mutations in sma-9, as in DBL-1 pathway genes, cause Sma and Mab phenotypes. sma-9 expression overlaps with that of DBL-1 pathway components. Anti-SMA-9 antibody staining reveals that SMA-9 is present in many tissues, where it localizes to cell nuclei; this is consistent with its predicted transcriptional cofactor function. In contrast to available reports on Drosophila shn, we find that the sma-9 locus generates a complex array of transcripts through alternative splicing, which are predicted to encode isoforms with diverse splicing, which are predicted to encode isoforms with

Mapping
sma-9 was previously mapped to linkage group X (Savage-Dunn et al., 2003). This map position was refined using SNP markers. lon-2sma-9 double mutants (Sma) were crossed with the Hawaiian strain CB4856. From these heterozygotes, Lon (Lon-2 non-Sma-9) recombinant progeny were selected. These progeny were tested for the presence of CB4856 SNP markers on the X-linked cosmids C36B7 (−2.04), Y49A10A (+1.91) and F11A1 (+2.2). The results demonstrate that sma-9 maps to the right of cosmid Y49A10A and within 0.1 map units of F11A1.

Body size measurements
Measurement of worm length, pharynx length and seam cell size was performed as described (Savage-Dunn et al., 2000; Wang et al., 2002).

Transgenic animals
The plasmid or cosmid DNA was microinjected into the gonadal syncytia of hermaphrodites, with rol-6 as a marker (Mello et al., 1991).

Sequencing sma-9 mutants
15 mutant animals were picked into 10 μl lysis buffer (10 mM Tris, pH 8.0; 50 mM KCl; 2.5 mM MgCl2; 0.45% Tween 20; 0.01% gelatin; 60 μg/ml proteinase K) and placed at −80°C for one hour. The frozen solution was heated to 60°C for 1 hour and then to 95°C for 20 minutes to generate crude lysate. PCR was carried out on mutant and wild-type genomic DNA templates, using platinum Taq and platinum Pfx mixtures as the DNA polymerases, and primers within genomic sequence. For wk55, the region from 5352 to 14827 in T05A10, covering the whole open reading frame (ORF), was sequenced. For qrc3, the regions from 5532 to 7558 and 11151 to 13081 in T05A10 were sequenced. The PCR fragments were sequenced directly, and all mutation sites were confirmed using a second primer.

Molecular cloning and sequencing
A total of 17 yk cDNA clones were sequenced in order to identify the structures of sma-9 transcripts. These yk clones were: yk128a11, yk128a2, yk1136g02, yk1134o06, yk856b10, yk1057d6, yk1216e10, yk1237d01, yk1103h10, yk127d10, yk1264e07, yk3289c and yk228h6. All yk cDNAs were gifts from Y. Kohara. Details for the primers used for sequencing the yk clones are available upon request. GenBank Accession Numbers for yk clones sequenced are AY390537-AY390553.

LiC RNA preparation from wild-type animals was performed as described at by Lui et al. (Lui et al., 1995). RT-PCR was performed using SUPERSCRIPT™ One-Step RT-PCR (Invitrogen). pCS234 was amplified using primers 5'-AAGCGGCCGCTAGACCATC-AGGCTACATTG-3' and 5'-AAGATCCGGTCTCAGGTTTGGTCAC-3', and cloned into pBluescript SK+ at NotI and BamHI. pCS234 starts from the predicted exon 1 (5' variant A in Fig. 3), splices out exon 4 and exon 5, and ends at exon 9, overlapping with yk328c9. pCS272 was amplified using primers 5'-AAGGACGAGCAGGCTGCGTGAG-3' and 5'-AAGGATCTGATGTTGCTCAGTG-3', and cloned into pBluescript SK+ at Smal and BamHI. pCS272 starts from exon 4 and ends at exon 9 overlapping both yk1285a11 and yk328c9. GenBank Accession Numbers for RT-PCR clone sequences are AY389809 and AY389810.

sma-9 upstream regions cloned by PCR were fused with the GFP reporter gene pPD117.01 (a gift from A. Fire) (Fig. 4I). Primers used for cloning were:

2.8 kb, 5'-GCCGCGGCGCAACACATTATTGTAAGTTG-3' and 5'-AAGGTATCATGAAGCTCGTCA-3';
5.5 kb, 5’-AAGCGGCGCAGTTCACTACATTGTATG-3’ and 5’-AAGGTACTTTCGCAATCTCTTAAAACACT-3’; 8.0 kb, 5’-AAGCGGCGCAGTTCACTACATTGTATG-3’ and 5’-AAGGTACTTTCGCAATCTCTTAAAACACT-3’; 1.5 kb, 5’-GCCGGCGAGTTCACTACATTGTATG-3’ and 5’-GCCGGCGCAGTTCACTACATTGTATG-3’; and 4.0 kb, 5’-GCCGGCAGTTCACTACATTGTATG-3’ and 5’-GCCGGCGCAGTTCACTACATTGTATG-3’.

The 2.8 kb fragment was cloned into pPD117.01 at NotI and KpnI (pCS231), then into pBlueScript SK+ (BanHI and KpmI), and finally into pPD117.01 at XbaI and KpmI (pCS251). The 5.5 kb fragment was cloned into pBlueScript SK+ (NotI and KpnI), then into pPD117.01 at SacII and KpnI (pCS252). The 8.0 kb fragment was cloned into pBlueScript SK+ (NotI and KpnI), then into pPD117.01 at SacII and KpnI (pCS253). The 1.5 kb fragment was cloned into pPD117.01 at SacII and NotI (pCS255). The 4.0 kb fragment was cloned into pPD117.01 at SacII and NotI sites (pCS256).

dsRNA

The templates used were the yk128a12 cDNA clone [containing sequences from predicted exons 1-7 and alternative exon I (Fig. 3)], and the yk228b6 cDNA clone [containing sequences from predicted exons 21-25, which are present in all 3′ variants as either translated (Class I isoforms) or untranslated (Class II and III isoforms) sequences]. yk128a12 was digested by EcoRI and KpnI; yk228b6 was digested by SmaI and KpnI. Then the digested DNA was extracted by phenol:chloroform once, precipitated by ethanol, dried in air, and dissolved in TE. 1 μg of the cut DNA was used to synthesize RNA by Stratagene RNA Transcription Kit. After that, the reaction solution was treated by RNase-free DNase1 at 37°C for 15 minutes. Then, the ssRNA was combined and extracted by phenol:chloroform once, precipitated by ethanol, dried in air and dissolved in 10 μl TE. The purified RNA was incubated at 68°C for 10 minutes and then at 37°C for 30 minutes. The dsRNA was microinjected directly into N2 animals without further treatment. Males with small body size were picked to score the Mab phenotype.

Antibodies and immunostaining

Anti-SMA-9 antibodies were generated against the 70 amino acids that are unique to the abundantly expressed class II isoforms. Plasmid pJKL5471.1, which contains the cDNA fragment corresponding to the 70 amino acids cloned into pGEX-2T, was transformed into BL21 cells. Fusion proteins were first purified using Glutathione sepharose 4B beads (Amersham Biosciences) and further purified by SDS-PAGE. Gel slices containing the purified fusion proteins were used to immunize rats by Cocalico Biologicals, PA. The resulting antibodies were tested by western blot analysis using bacteria generated fusion proteins. For immunostaining, wild-type N2 and w;5 mutant animals were fixed following the protocol of Hurd and Kemphues (Hurd and Kemphues, 2003). Rat anti-SMA-9 isofrom II antisera (serum bxEx46) was used at 1:2000 dilution. Affinity-purified Cy3-conjugated donkey anti-rat secondary antibodies (Jackson Immunoresearch Laboratories) were used at 1:100 dilution.

Ray neuron fate expression

A- and B-type ray neuron generation in sma-9, DBL-1 pathway mutant or wild-type males was assessed by examining the expression of the A- and B-type neuron marker OSM-6::GFP, and the B-type neuron marker PKD-2::GFP (Collett et al., 1998; Barr and Sternberg, 1999) (L. Jia and S.W.E., unpublished). In sma-9 and in DBL-1 pathway mutants, the expression patterns of these reporters did not differ significantly from wild type indicating that the absence of dopaminergic or serotonergic marker expression from certain ray neurons was not due to loss of the affected neuron (data not shown). The presence of dopamine in male ray neurons was assessed using formaldehyde-induced fluorescence (FIF) as described by Lints and Emmons (Lints and Emmon, 1999). FIF patterns were found to be consistent with the expression pattern of the dopaminergic marker CAT-2::GFP in strains examined. Serotonergic fate expression was assessed by staining the whole animal with anti-serotonin antisera as described by Loer and Kenyon (Loer and Kenyon, 1993). To assist with B-type ray neuron identification, strains examined also carried a pkd-2::gfp reporter array (bxEx14) (Barr and Sternberg, 1999) (L. Jia and S.W.E., unpublished). SHT-antibody staining patterns were found to be consistent with the expression pattern of the serotonergic fate marker TPH-1::GFP.

Heat shock and temperature shifts

Heat shock DBL-1 experiments were performed as described (Lints and Emmons, 1999). Individual sma-9 or wild-type males carrying hs::dbl-1::cat-2::gfp arrays bxEx46, or bxEx47 or control empty vector/cat-2::gfp arrays bxEx44 or bxEx45, were staged by examination of seam cells with Nomarski optics (Sulston and Horvitz, 1977). Males of the Rn stage were transferred to a siliconized Eppendorf tube containing 100 μl of M9 buffer (Brenner, 1974), which was placed in a 30°C circulating water bath for 30 minutes. After heat-shock, animals were recovered, placed at 20°C on pre-equilibrated OP50-seeded plates, allowed to develop to adulthood and then scored for CAT-2::GFP expression in the rays.

Results

sma-9 is a new component of the DBL-1 pathway

In C. elegans, activity of the DBL-1 pathway during development regulates body size and the identity of at least two male-specific copulatory structures, the male sensory rays and the spicules (Savage et al., 1996; Lints and Emmons, 1999). The hypodermis is the crucial DBL-1-responsive tissue for body size regulation (Wang et al., 2002). The hypodermis, most of which is made up of a single multinucleate syncytium, hyp7, surrounds the animal and secretes the cuticle. During each larval stage, two lateral rows of hypodermal seam cells divide in a stem-cell-type lineage with one daughter cell fusing into hyp7 for growth. In the males, posterior cells of the lateral seam execute a sex-specific pattern of cell division, generating nine bilateral pairs of male-specific sensory rays. The DBL-1 pathway is required for the identities of rays 5, 7 and 9, with mutations resulting in sensory ray fusions and alterations in the expression of ray neuron neurotransmitters.

To identify additional pathway components, and in particular those that confer specific responses to pathway activity, we performed forward genetic screens for mutants affected in body size (Savage-Dunn et al., 2003) and male ray neurotransmitter identity (Lints and Emmons, 2002). A total of five recessive mutant alleles of a newly defined locus, sma-9, were isolated in these screens (wk and bx alleles). As only viable mutants were selected in these screens, we wished to determine whether null alleles of sma-9 cause more severe defects or lethality. Because sma-9(wk62)/Df animals are viable, a screen for mutations that fail to complement sma-9(wk62) should uncover more severe or homozygous lethal alleles of sma-9 if they can be created. An additional nine sma-9 alleles were identified in such a screen (see Materials and methods), but none of them showed lethality or defects more severe than the previously isolated alleles. Thus, these alleles show the full range of defects associated with sma-9 loss of function. sma-9(d9) mutants, similar to DBL-1 pathway mutants, have a small body size, an abnormal sensory ray identity (ray 8-9 fusions) and crumpled spicules (Fig. 1, Table 1), suggesting that sma-9 might define a new pathway.
component. Most of our analyses have been carried out using \textit{sma-9(wk55)} mutants, which display a strong mutant phenotype.

**sma-9 encodes a Shn-like protein**

We mapped \textit{sma-9} by positional cloning using standard genetic markers and SNPs. The mapping data placed \textit{sma-9} on linkage group \textit{X}, to the right of cosmid Y49A10A (+1.91) and within 0.1 map units of F11A1 (+2.2) (see Materials and methods). We microinjected cosmids of this region into a \textit{sma-9(wk55)} (lf) mutant background and assessed their ability to rescue body size and male tail defects. Cosmid T05A10 rescued \textit{sma-9} mutant phenotypes, including body size defects, the male tail defects (ray 8-9 fusions were reduced from 49% to 23%), and crumpled spicules (Fig. 1, Table 1). Significantly, T05A10 (GenBank Accession Number Z68108) is predicted to encode a transcription factor homologous to \textit{Drosophila} Shn, a Smad cofactor in the Dpp pathway. We hypothesized that T05A10.1 corresponds to \textit{sma-9}. Sequencing of the corresponding ORF in \textit{sma-9(wk55)} and \textit{sma-9(qc3)} genomic DNA identified nonsense mutations in each (Fig. 2A): \textit{wk55} converts Arg1163 to a stop codon and \textit{qc3} converts Gln204 to a stop codon. The nonsense mutations in \textit{wk55} and \textit{qc3} are predicted to terminate most, but not all, isoforms identified in the cDNA sequences discussed below, suggesting that they are likely to be strong loss-of-function, but not null, alleles. Furthermore, disruption of T05A10.1 function by dsRNAi using 3¢ cDNA sequences, or combined 5¢ and 3¢ cDNA sequences (see Materials and methods), phenocopies \textit{sma-9} mutants, causing small body size in both sexes, and fusion between male rays 8 and 9 but no additional phenotypes (Fig. 1, Table 1). Therefore the predicted gene T05A10.1 corresponds to \textit{sma-9}.

The longest conceptual SMA-9 sequence, based on the isolated cDNA clones discussed below, is 2170 amino acids (aa) in length and contains a Gln-rich N terminus, including several repeats of a QQQQL sequence of unknown function, and seven C2H2 zinc finger motifs at the C terminus clustered into two pairs and one triplet (Fig. 2A). The first pair of zinc fingers is located in the middle of the sequence, the second pair is near the C terminus and the triplet is between these two pairs. Like Shn, SMA-9 is rich in Ser and Thr. There are two acidic-residue-rich domains (ARD) that may correspond to transcriptional activation domains, one N-terminal to the zinc finger region and the other following the first pair of zinc fingers. The whole sequence contains four predicted NLSs, three at the N terminus and one at the C terminus, consistent with a function in the nucleus. Five S/TPKK motifs surround the zinc finger regions; these motifs are putative DNA-binding domains and may be regulated by phosphorylation (Hill et al., 1990).

A similarity search of GenBank revealed that \textit{sma-9} shares high sequence homology with a zinc finger transcription factor family that includes \textit{Drosophila} Shn (BLAST E value=4e-24) (Arora et al., 1995) and vertebrate Shn1 family members, human major histocompatibility complex-binding protein 1 (MBP1)/PRDII-BF1 (E value=2e-17) (van’t Veer et al., 1992; Fan and Maniatis, 1990) and mouse \(\alpha\)A-crystallin-binding protein 1 (\(\alpha\)A-CRYBP1; E value=8e-18) (Nakamura et al., 1990). Similarities among them include the presence of multiple zinc fingers, NLS, ARD and S/TPKK motifs, as well as stretches of sequences rich in Gln and in Ser/Thr. In SMA-9, the first pair of zinc fingers has 77% identity to the second pair in Shn, 76% to the second pair in MBP1/PRDII-BF1 and 74% to the second pair in \(\alpha\)A-CRYBP1 (Fig. 2B). Therefore, SMA-9, MBP1, \(\alpha\)A-CRYBP1 and Shn may derive from a common ancestral gene, and this pair of zinc fingers may...
contribute to a conserved role for Shn proteins. The SMA-9 triplet of zinc fingers has 45% identity to the Shn triplet. This domain is absent from MBP1 and A-CRYBP1 (Fig. 2B), suggesting its elimination during vertebrate evolution or its acquisition in the fly-worm lineage. The SMA-9 second pair of zinc fingers has no similarity to the other family members (Fig. 2C), indicating a unique function in C. elegans. An alternative 70 aa C terminus is also unique to C. elegans.

sma-9 displays complex alternative splicing

To verify the sma-9 ORF predicted by Genefinder, we isolated sma-9 cDNAs by RT-PCR and analyzed clones available from cDNA libraries. A total of 17 yk cDNA clones (see Materials and methods) were sequenced; one other cDNA clone (G-dvpl1458.x) had been previously identified (Walhout et al., 2000). RT-PCR was performed to analyze the transcript structure at the 5' end of the gene (see Materials and methods). With the exception of yk1285a11 and yk1237d01 discussed below, none of the cDNA clones are full length, so we cannot be sure which 5' and 3' variants are present in contiguous transcripts in vivo.

Analysis of sma-9 cDNA clones showed complex alternative splicing at both 5' and 3' ends (Fig. 3). At least three 5' end forms were detected. Class A (represented by pCS234) contains the predicted initiation ATG (predicted exon 1), but lacks predicted exons 4 and 5 where the first NLS is located. Class B (represented by yk1285a11) contains an alternative upstream exon (alternative exon 1), splices out part of predicted exon 1 including the ATG, and lacks part of exon 4 but contains the first NLS in exon 5. In addition, this transcript is trans-spliced (SL1) and terminates with a poly(A) tail after exon 7.

To test whether all exon 4 containing transcripts terminate in exon 7, we performed RT-PCR using primers in exon 4 and exon 9, and generated the product pCS272 (see Materials and methods), representing Class C.

The 3' end of sma-9 is even more diversified than the 5' end. Alternative splicing exists in predicted exons 15, 20 and 21 that would result in variable numbers of zinc finger clusters being expressed in different isoforms (Fig. 3) and in different C termini, including a 70 aa sequence that is unique for C. elegans (Fig. 2A). In Class I isoforms (5 cDNA clones), all of the zinc finger motifs are present; in Class II isoforms (10 cDNA clones), the second pair is missing and the unique 70 aa sequence is present; and in Class III isoforms (1 cDNA clone), only the first pair of zinc fingers is translated. Based on the numbers of cDNA clones isolated, Class II isoforms are predicted to be most abundant. Thus, as a result of alternative splicing, SMA-9 isoforms would differ in the numbers of NLSs, zinc finger motifs and S/TPKK motifs, which could allow differences in subcellular localization, transcriptional activity, DNA binding ability or expression pattern (discussed below).

Interestingly, in one Class IIb cDNA clone, yk1237d01, exon 11 is trans-spliced to SL2, the trans-spliced leader sequence associated with downstream genes in polycistronic operons (Blumenthal et al., 2002). The SL2-spliced transcript may therefore form an operon with the upstream transcript defined by yk1285a11. To our knowledge, this is the first published report of competing cis- and SL2 trans-splicing to the same splice acceptor sequence. However, because these cDNAs were rare they might represent low abundance messages, messages with cell-type specific expression patterns or spurious events.

### Table 1. Mab phenotypes of sma-9 strains

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Percentage fusion of rays 4-5</th>
<th>Percentage fusion of rays 6-7</th>
<th>Percentage fusion of rays 8-9</th>
<th>n (sides)</th>
<th>Percentage crumpled spicules</th>
<th>n (tails)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>74</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Single mutants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sma-9(wk55)</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td>105</td>
<td>46</td>
<td>50</td>
</tr>
<tr>
<td>sma-9(wk62)</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>102</td>
<td>66</td>
<td>49</td>
</tr>
<tr>
<td>sma-9(wk71)</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>104</td>
<td>32</td>
<td>52</td>
</tr>
<tr>
<td>sma-9(wk82)</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>103</td>
<td>43</td>
<td>51</td>
</tr>
<tr>
<td>sma-9(qc1)</td>
<td>0</td>
<td>0</td>
<td>67</td>
<td>30</td>
<td>32</td>
<td>20</td>
</tr>
<tr>
<td>sma-9(qc3)</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>100</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td>sma-9(qc5)</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>59</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>sma-9(qc6)</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>80</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>sma-9(qc7)</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>112</td>
<td>53</td>
<td>80</td>
</tr>
<tr>
<td>sma-9(qc8)</td>
<td>0</td>
<td>0</td>
<td>43</td>
<td>86</td>
<td>63</td>
<td>47</td>
</tr>
<tr>
<td>sma-9(qc9)</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>93</td>
<td>58</td>
<td>34</td>
</tr>
<tr>
<td>sma-9(qc10)</td>
<td>0</td>
<td>0</td>
<td>47</td>
<td>51</td>
<td>88</td>
<td>16</td>
</tr>
<tr>
<td>sma-9(qc11)</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>105</td>
<td>51</td>
<td>88</td>
</tr>
<tr>
<td>sma-3(wk30)</td>
<td>26</td>
<td>64</td>
<td>21</td>
<td>121</td>
<td>100</td>
<td>30</td>
</tr>
<tr>
<td>Double mutants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sma-3(wk30);sma-9(wk55)</td>
<td>11</td>
<td>40</td>
<td>16</td>
<td>81</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>sma-3(wk30);sma-9(wk62)</td>
<td>30</td>
<td>45</td>
<td>21</td>
<td>100</td>
<td>100</td>
<td>30</td>
</tr>
<tr>
<td>sma-9 rescue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sma-9(wk55);qcEx49</td>
<td>0</td>
<td>1</td>
<td>23</td>
<td>79</td>
<td>9</td>
<td>42</td>
</tr>
<tr>
<td>dsRNAi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exons 1-7</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>102</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>exons 21-25</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>70</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td>exons 1-7 and 21-25</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>75</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

All strains contain him-5(e1490). 
qcEx49 contains cosmid T05A10 and plasmid pRF4.
ND, not determined.
with no functional significance. The region between exons 7 and 11 has some features of an intercistronic region, but is clearly not typical. Forty-eight nucleotides downstream of exon 7 is an imperfect match to the AAUAAA sequence necessary for polyadenylation: AAUUAAA. Upstream of exon 11 is a U-rich region, UUAUCUCUGUGUAUAAU, reminiscent of identified sequences that are necessary for SL2-mediated trans-splicing (Huang et al., 2001). However, the distance between the two transcripts is 2.3 kb, whereas a typical intercistronic region is between 100 and 120 bp (Blumethal et al., 2002). The results of RNAi suggest that these trans-spliced messages encode isoforms with distinct but overlapping functions. Inactivation of exons 1-7 results in a wild-type male tail phenotype, whereas inactivation of exons 21-25, which would target the \( yk1237d01 \) IIb isoform, results in significant frequencies of ray 8-9 fusions (Table 1). Conversely, both experiments resulted in a similar Sm9 body size (Fig. 1 and data not shown). The strong male tail phenotype caused by \( sma-9(wk55) \), which should not disrupt either of these isoforms, could be due either to instability of the transcript containing a premature termination codon, or to the expression of a truncated protein with antimorphic properties. Significantly, no expression of Class II isoforms was detected in \( sma-9(wk55) \) by immunohistochemistry (see below).

**sma-9 is widely expressed**

To understand \( sma-9 \) function during development, we examined the \( sma-9 \) transcriptional expression pattern (by fusing \( sma-9 \) upstream promoter regions with a GFP reporter gene) and protein localization (by immunohistochemistry using anti-\( SMA-9 \) antibodies) (Fig. 4). \( sma-9 \) promoter-driven GFP was expressed in the ventral nerve cord (VNC; Fig. 4A), pharynx and intestine (Fig. 4B), seam cells (Fig. 4C), excretory canal, vulva and spermatheca (data not shown). Expression was observed from L1 to adult stages but not during embryonic development. Significantly, expression in the lateral seam coincided with the crucial period for \( sma-9 \) activity in body size regulation (see below). \( sma-9 \) promoter-driven GFP was detected in the lateral seam only from the L1 to the L3 stages, and not in the L4 stage or in the adult.

![Fig. 2. sma-9 encodes a zinc finger transcription factor.](image-url)
Fig. 3. Alternative splicing of sma-9 transcripts. sma-9 cDNA clones reveal a complex alternative splicing at both 5' and 3' ends. At least three N terminus and seven C terminus forms were found. Predicted exons 4, 15 and 20 are labeled. Approximate locations of the termination codons in qa3 and wk55 are shown above the predicted exon structure. Gray boxes represent unique C-terminal sequences; an alternative exon 15 with two extra bases is indicated in yellow. cDNA clones representing each variant are: class A – pCS234 from RT-PCR; class B – yk128a8; class C – pCS272 from RT-PCR; class Ia – yk128a8, yk1136g02 and yk1109f01; class Ib – yk43h3; class Ic – G_dvp11458.x (Walhout et al., 2000); class IIa – yk6d10, yk86d1c and yk1134e06; class Ilb – yk1237d01, which is SL2-spliced, yk568b10, yk1057a6, yk1216e10, yk1103h10 and yk127d10, which is also missing the intron between exon 21 and 22 (not shown); class IIc – yk1264e07; and class III – yk328e9. Only the trans-spliced cDNA clones yk1285a11 and yk1237d01 appear to be full length.

Fig. 4. sma-9 is widely expressed. sma-9 upstream sequences (K) were fused with a GFP reporter gene. (A-C) GFP expression in wild-type transgenic animals. The transgenic animals display a strong fluorescence in the VNC, pharynx and intestine, from stages L1 (not shown) to adult (A,B). L2 stage (C) animals show expression in the seam cells that disappears after L3 (Nomarski image in D). Arrows in C and D indicate seam cell nuclei. (E,G,I) Immunostaining by anti-SMA-9 antibodies against the unique 70 aa C terminus present in class II isoforms. (F,H,J) DAPI staining of the same animals. (E,F) N2; (G,H) sma-9(wk55); (I,J) enlarged view of the anterior region of a wild-type late L1 stage animal. In all animals, anterior is to the left, arrowheads show pharyngeal nuclei, black arrows show intestinal nuclei and white arrows show hypodermal nuclei. (K) Different sma-9 upstream regions used to construct GFP reporters.
Endogenous SMA-9 protein localization was examined by immunohistochemistry using antibodies against the 70 aa C. elegans-specific domain present in the abundant Class II isoforms (see Materials and methods). Antibody staining detects SMA-9 protein in most, if not all, somatic nuclei of wild-type animals, but not in sma-9(wk55) animals (Fig. 4E,G). The lack of protein expression in sma-9(wk55) mutants supports our conclusion that this allele represents a strong loss of function allele. The localization of SMA-9 to the nucleus is consistent with a function as a transcriptional cofactor. Furthermore, the immunolocalization confirms SMA-9 expression in the pharynx (arrowheads), intestine (black arrows), hypodermis (white arrows) and VNC (Fig. 4I,J). Therefore, the expression pattern of SMA-9 overlaps with that of DVL-1 pathway components (Savage-Dunn et al., 2000; Krishna et al., 1999; Gunther et al., 2000; Suzuki et al., 1999).

Interestingly, different promoter regions do not give rise to the same expression pattern. The 2.8 kb, 5.5 kb and 8.0 kb constructs (upstream of predicted exon 1; Fig. 4K) show strong fluorescence in all detected tissues except the seam cells. The 1.5 kb construct (−5500 bp to −4000 bp relative to predicted exon 1) displays expression in the seam cells, VNC and excretory canal only. The 4.0 kb construct (−8000 bp to −4000 bp) generates the complete expression pattern, indicating the presence of redundant transcriptional elements. The 5.5 kb and 8.0 kb fragments are identical to the 1.5 kb and 4.0 kb fragments, respectively, except for the addition of sequences from −4000 bp to −1 bp that include the alternative exon 1 (Fig. 4K). The addition of these sequences abolishes expression in the seam cells in these reporters. These results could be due to the presence of a seam-cell specific repressor element between −4000 bp and −1 bp. Alternatively, transcription and splicing in the seam cells might specifically generate the Class B variant that initiates translation in alternative exon 1, which would render the GFP sequences in the 5.5 kb and 8.0 kb constructs out of frame.

sma-9 functions downstream of DVL-1 to regulate body size

The DVL-1 pathway regulates body size throughout post-embryonic development (Savage-Dunn et al., 2000). To understand sma-9 function in body size development, we measured worm length, pharynx length and the seam cell size of DVL-1 pathway mutants and sma-9 mutants at various times after embryogenesis. The sma-9 hatched L1 larva is indistinguishable from both wild type and the DVL-1 pathway mutants (Fig. 5A). In L2 and L3 stages, sma-9 animals have the same size and growth rate as the DVL-1 pathway mutants have. Furthermore, the sma-9(lf) pharynx length and seam cell length is indistinguishable from that of DVL-1 and sma-9. Data for N2 and DVL-1 is from Wang et al. (Wang et al., 2002). (C) Model of sma-9 function in the DVL-1 pathway. In body size development, sma-9 functions in early larval stages and may be replaced by other transcriptional cofactors in late larval stages. In male tail development, sma-9 prevents fusions in rays 8-9 specifically, but regulates cat-2 activity, the rate-limiting step in dopamine expression, in all rays. Other cofactors may be involved in ray 4-5 and 6-7 fusions. For spicule development, both sma-9 and other cofactors are probably required.
sma-9 is required only in early larval development and is dispensable later.

To better understand the relationship between sma-9 and the DBL-1 pathway, we analyzed the effect of sma-9(lf) on dbl-1 overexpression and lon-1(lf) phenotypes (Fig. 1). Overexpression of the ligand gene dbl-1 from an integrated array in a wild-type background gives the animal a Lon phenotype (Suzuki et al., 1999) (Fig. 1J). However, in the sma-9(wk55) (lf) background the Lon phenotype is no longer observed and the animals are Sma (Fig. 1K). This suggests that sma-9 functions genetically downstream of dbl-1 and is required for regulating body size. lon-1 acts downstream of the DBL-1 pathway and encodes a cysteine-rich secretory protein (CRISP) (Maduzia et al., 2002; Morita et al., 2002). It has been shown that lon-1 is negatively regulated by the pathway, and that the Lon phenotype is expressed in late larval stages and in adulthood. Unlike lon-1, DBL-1 pathway mutants that are Lon, lon-1; sma-9 doubles, are neither Lon nor Sma but are of wild-type size in adulthood (Fig. 1G). This suggests that sma-9 acts independently of lon-1, consistent with their functions in different larval stages.

**sma-9 functions downstream of DBL-1 in regulation of male sensory ray identity**

The male tail bears nine bilateral pairs of sensory rays (Fig. 1A). Each ray contains two neurons, an A- and a B-type neuron, and a structural cell. The cells of a single ray derive from a common precursor cell Rn (n stands for rays 1 to 9) generated by the posterior seam at L3. The DBL-1 pathway regulates multiple aspects of male ray identity and morphogenesis of the spicules. In DBL-1 pathway mutants, rays 5, 7 and 9 often fuse with their anterior neighbor (Savage et al., 1996; Morita et al., 1999; Krishna et al., 1999), and the fate of neurons within these rays is altered. The A-type neurons of rays 5, 7 and 9 (R5A, R7A and R9A) express a dopaminergic (DA) fate with reduced frequency compared with wild type (Lints and Emmons, 1999), and in the B-type neurons of rays 5 and 9 (R5B and R9B), serotonergic fate is ectopically expressed or abolished, respectively (R.L. and S.W.E., unpublished) (Table 2). Reduced DA expression is associated with reduced expression of the dopaminergic marker CAT-2::GFP (see Materials and methods), suggesting that cat-2 may be a direct or indirect target of the DBL-1 pathway. In addition to having ray defects, the spicules of DBL-1 pathway mutants fail to elongate and are frequently crumpled.

In sma-9(lf) mutants, ray 8-9 fusions occur frequently (Fig. 1, Table 1), but ray 4-5 and ray 6-7 fusions do not. Ray neurons R5A, R7A and R9A express a DA fate with reduced frequency compared with wild type; however, the distribution of frequency and the intensity of DA fate marker expression differ from that of DBL-1 pathway mutants (Table 2). Also in contrast to DBL-1 pathway mutants, serotonergic patterning in sma-9 mutants is largely similar to wild type, except for a moderate reduction in the frequency of serotonergic fate expression in R9B (reduced from 100% in wild type to 75% in sma-9 mutants) and weak ectopic expression of this fate in R5B and R7B (in less than 5% of sides scored for each ray neuron) (Table 2). sma-9 mutants also exhibit crumpled spicules, although the defect is less severe and occurs with lower penetrance than in DBL-1 pathway mutants (Table 1). Thus, as for body size, sma-9(lf) mutant phenotypes in the male tail overlap considerably with, but are not identical to, those of DBL-1 pathway mutants.

Two observations support the notion that in male ray patterning sma-9 acts downstream of the DBL-1 pathway. First, the sma-9; DBL-1 pathway double mutant phenotype was similar to that of the DBL-1 pathway single mutants. Rays 5, 7 and 9 fuse with their anterior neighbor (Table 1), and serotonergic fate is expressed in ray neurons R1B, R3B and R5B in all animals but not in R9B (Table 2). Second, in a wild-type background, induction of a heat-shock promoter driven dbl-1 transgene (HS::dbl-1) during the Rn stage of ray development causes ectopic expression of the DA marker CAT-2::GFP in R3A, R4A, R6A and R8A (Table 3). By contrast, in a sma-9 mutant background, HS::dbl-1 gene activation does not induce ectopic expression of CAT-2::GFP and animals display a sma-9 mutant phenotype (Table 3). Together, these experiments suggest that sma-9 functions genetically downstream of DBL-1 signaling, and that its activity is necessary for mediating the effects of the DBL-1 pathway in male ray patterning. In addition, the HS::dbl-1 experiments reveal that, like DBL-1 pathway components, sma-9 activity is not restricted to rays 5, 7 and 9, and that the gene can function in other rays.

**Discussion**

**sma-9 functions in the DBL-1 pathway**

We have demonstrated that SMA-9 belongs to a transcription factor family that includes *Drosophila* Shn, human MBP1/PDII-BF1 and mouse αA-CRYBP1. Based on the similarity of the phenotypes of sma-9(lf) and DBL-1 pathway...
Table 3. Percentage frequency of dopaminergic (DA) fate expression in male ray A-type neurons in response to heat-shock induction of a dbl-1 transgene (HS::dbl-1)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Treatment</th>
<th>Ray</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>no heat-shock</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>100</td>
<td>–</td>
<td>100</td>
<td>–</td>
<td>100</td>
<td>–</td>
<td>100</td>
</tr>
<tr>
<td>+ heat-shock</td>
<td>–</td>
<td>–</td>
<td>60</td>
<td>100</td>
<td>20</td>
<td>100</td>
<td>20</td>
<td>100</td>
<td>–</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>sma-9 (wk55)</td>
<td>no heat-shock</td>
<td>–</td>
<td>–</td>
<td>10*</td>
<td>10*</td>
<td>–</td>
<td>10*</td>
<td>15*</td>
<td>–</td>
<td>15*</td>
<td></td>
</tr>
<tr>
<td>+ heat-shock</td>
<td>–</td>
<td>–</td>
<td>10*</td>
<td>5*</td>
<td>10*</td>
<td>15*</td>
<td>–</td>
<td>18*</td>
<td>–</td>
<td>18*</td>
<td></td>
</tr>
</tbody>
</table>

The DBL-1 ligand was provided from a ubiquitously expressed, heat-shock inducible dbl-1 transgene (HS::dbl-1) in the genetic background indicated (see Materials and methods).

Forty sides were scored per treatment/genotype. See legend for Table 2.

In the Mab phenotype, sma-9 loss of function appears to have more restricted effects than the loss of other DBL-1 pathway components does. sma-9 mutants display ray 8-9 fusions at high frequency, but never ray 4-5 or ray 6-7 fusions (Table 1). This specificity contrasts with the function of sma-9 in regulating neurotransmitter expression in the same lineages: sma-9, like dbl-1, is required for patterning neurons within rays 5 and 7, as well as in ray 9. Therefore, the sma-9(/f) mutant phenotype is not merely a weaker version of the DBL-1 pathway loss-of-function phenotype, but rather a more specific one. For most aspects of the phenotype that are in common, the sma-9 defects are no less severe than those of DBL-1 pathway mutants. These results lead us to propose the model presented in Fig. 5C, in which target gene specificity is determined in part by the activity of sma-9 and in part by other transcriptional cofactors.

Alternative splicing may produce SMA-9 isoforms with different activities

The presence of various SMA-9 isoforms suggests multiple functions in signaling. One purpose of the complex splicing pattern may be to generate isoforms with different subcellular localizations. Vertebrate Shn isoforms KRC (Oukka et al., 2002) and GAAP-1 (Lallemand et al., 2002) have been demonstrated to reside in both the nucleus and the cytoplasm, suggesting that their activity might be regulated by subcellular compartmentalization. In the case of SMA-9, the localization properties of predicted isoforms are not yet known. However, SMA-9 isoforms differ in the number of NLSs. N-terminal variant A is predicted to lack the strongest NLS encoded by exon 5 (Fig. 3). More strikingly, the SL2-spliced variant of isoform IIb would lack all of the predicted NLS sequences, suggesting either a function in the cytoplasm or a need to interact with a nuclear-localized factor to shuttle into the nucleus. A second purpose of alternative splicing may be the generation of cell-type specific isoforms. The use of diverse regulatory sequences both upstream and downstream of the alternative exon 1 could contribute to such cell-type specificity. We suggest that various SMA-9 isoforms have distinct but overlapping functions in the DBL-1 pathway. Finally, some SMA-9 isoforms may not function in the DBL-1 pathway but may serve other functions. Interestingly, sma-9(/f) mutants have a mesodermal defect that is independent of DBL-1 function (M.L.F. and J. Liu, unpublished).

Conservation of Shn/SMA-9 function in BMP signaling

In Drosophila, Shn has been identified as a Smad cofactor in the Dpp BMP-related pathway (Dai et al., 2000). The identification of SMA-9, a Shn homolog, in a BMP signaling pathway in a distantly related animal phylum provides the first evidence for a conserved role for Shn proteins in BMP signaling. Drosophila shn mutants have a less severe phenotype than dpp null mutants (Arora et al., 1995). Similarly, sma-9 mutants at first appear to display weaker defects in both body size and male tail pattern than DBL-1 pathway mutants do. However, because sma-9 mutants are viable, we have been able to analyze phenotypes throughout the course of development, and we make the assessment that the sma-9 mutant phenotype is not weaker but rather is more specific than that of the DBL-1 pathway mutants as discussed above. These

...
results suggest that Shn may also function in a specific manner in the Drosophila Dpp pathway.

Vertebrate homologs of SMA-9 and Shn have been identified, but have not been shown to function in BMP signaling. Our results on SMA-9 suggest that it may be necessary to consider the existence of multiple isoforms with divergent functions. One motif that may be associated with a conserved function is the first pair of zinc fingers in SMA-9, which is highly conserved among these proteins (Fig. 2C). Other motifs may be associated with unique functions. It is likely that the complexity of functions mediated by this family is only just beginning to be appreciated.

We thank Lisa Madzuza and Richard Padgett for the BW1940 cts40 X strain; Yuji Kohara for cDNA clones; Yuval Hiltzik for assistance with the non-complementation screen; Thomas Blumenthal for helpful discussions on trans-splicing; and Daryll Hur and Ken Kemphues for suggestions on immunohistochemical staining. M.L.F. and J. Liu sequenced all the sma-9 yk cDNA clones, generated the anti-SMA-9 antibodies and performed the immunohistochemical staining. This research was supported by RGP 98-230-01-DDC from the American Cancer Society and a PSC-CUNY research award to C.S.S.; by NIH R01 GM06953 to J. Liu; and by NIH R01 NS38906 to S.W.E. M.L.F. was partially supported by NIH T32 GM07617; R.L. Staining. This research was supported by RPG 98-230-01-DDC from Kemphues for suggestions on immunohistochemical staining. M.L.F. helpful discussions on trans-splicing; and Daryl Hurd and Ken with the non-complementation screen; Thomas Blumenthal for

References


encoding a protein that binds to a cis sequence motif shared with the major histocompatibility complex class I and other genes. Mol. Cell. Biol. 10, 3700-3708.


