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

Jun Liang[1], Robyn Lints[2,3], Marisa L. Foehr[4], Rafal Tokarz[1], Ling Yu[1], Scott W. Emmons[2], Jun Liu[4] and Cathy Savage-Dunn[1,*]

[1]Department 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
[2]Department of Molecular Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA
[3]Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA
[4]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA
*Author for correspondence (e-mail: csavage@qc1.qc.edu)

Accepted 11 September 2003

Development 130, 6453-6464
Published by The Company of Biologists 2003
doi:10.1242/dev.00863

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; Lallemand 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 differences in the number of nuclear localization signals (NLS) and zinc finger motifs. The complexity of the *sma-9* locus suggests that different isoforms may have diverse functions and subcellular localizations. Further analysis of the *sma-9* mutant phenotype suggests that SMA-9 activity is spatially and temporally restricted relative to that of other DBL-1 pathway components. We propose that some SMA-9 isoforms function as Smad transcriptional cofactors to confer specific responses to DBL-1 pathway activation in regulating body size and male tail development.

Materials and methods

Strains

C. elegans strains were cultured using standard methods and grown at 20°C unless otherwise noted (Brenner, 1974). In addition to strains generated in this work, the following strains were used:

N2 (wild type);

LG III, *sma-3(wk30), sma-4(e729), lon-1(e185)*;

LG IV, *dbl-1(wk70)*;

LG V, *sma-1(e30), him-5(e1490), Is[tph-1::gfp + rol-6(su1006)]* (Sze et al., 2000);

bxIs14 an integrated derivative of *pkd-2::gfp*-containing array *syEx313* (Barr and Sternberg, 1999) (L. Jia and S.W.E., unpublished);

LG X, *sma-9(wk55, wk62, wk71, wk82)* (Savage-Dunn et al., 2003);

sma-9(bx120), lon-2(e678), dbl-1(ctIs40) (Suzuki et al., 1999);

cat-2::gfp complex extrachomosomal arrays *bxEx44, bxEx45, bxEx46 and bxEx47* (Lints and Emmons, 1999); and

mnIs17, an integrated derivative of *osm-6::gfp* array *mnEx64* (Collet et al., 1998).

The isolation of *wk* and *bx* alleles was described previously (Savage-Dunn et al., 2003; Lints and Emmons, 2002). The *qc* alleles were isolated in a non-complementation screen. N2 males were mutagenized with EMS (Brenner, 1974) and mated with *sma-9(wk62)unc-7(e5)* hermaphrodites. The F1 generation was screened for Sma non-Unc animals. From a screen of approximately 4000 F1 cross progeny, nine new alleles of *sma-9* were isolated.

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). 20 ng/μl cosmid DNA was injected into *sma-9(wk55)* for rescue. 10 ng/μl plasmid DNA of GFP constructs was injected into N2.

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 $MgCl_2$; 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* mixture as the DNA polymerase, and primers within genomic sequence. For *wk55*, the region from 5532 to 14827 in T05A10, covering the whole open reading frame (ORF), was sequenced. For *qc3*, 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: *yk1285a11*, *yk128a8*, *yk1136g02*, *yk1109f01*, *yk43h3*, *yk6d10*, *yk864c1*, *yk1134e06*, *yk856b10*, *yk1057a6*, *yk1216e10*, *yk1237d01*, *yk1103h10*, *yk127d10*, *yk1264e07*, *yk328c9* 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 clone sequences are AY390537-AY390553.

LiCl 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′-AAGCGGCCGCATGAGCCATC-AGGCAATTGG-3′ and 5′-AAGGATCCGGTTCAAGGTTTTGTG-TCAC-3′, and cloned into pBluescript SK+ at *Not*I and *Bam*HI. 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′-AACCCGGG-CCCCTCGCTCTCCAAA-3′ and 5′-AAGGATCCTGATGGTCCT-TG-3′, and cloned into pBluescript SK+ at *Sma*I and *Bam*HI. 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′-GCGCGGCCGCAAACATTTGTGAAGTTG-3′ and 5′-AAGGTACCTTCGCCAATTCTAAAACCACT-3′;

5.5 kb, 5′-AAGCGGCCGCGAGTTCACACAGTTTATGAT-3′ and 5′-AAGGTACCTTCGCCAATTCTAAAACCACT-3′;

8.0 kb, 5′-AAGCGGCCGCCCATCCAATATTCAATTCTT-3′ and 5′-AAGGTACCTTCGCCAATTCTAAAACCACT-3′;

1.5 kb, 5′-GGCCGCGGAGTTCACACAGTTTATGAT-3′ and 5′-GGGCGGCCGCCGAAAATTGCAGGTCTG-3′; and

4.0 kb, 5′-GGCCGCGGCCCATCCAATATTCAATTCTTTA-3′ and 5′-GGGCGGCCGCCGAAAATTGCAGGTCTG-3′.

The 2.8 kb fragment was cloned into pPD117.01 at *Not*I and *Kpn*I (pCS231), then into pBluescript SK+ (*Bam*HI and *Kpn*I), and finally into pPD117.01 at *Xba*I and *Kpn*I (pCS251). The 5.5 kb fragment was cloned into pBluescript SK+ (*Not*I and *Kpn*I), then into pPD117.01 at *Sac*II and *Kpn*I (pCS252). The 8.0 kb fragment was cloned into pBluescript SK+ (*Not*I and *Kpn*I), then into pPD117.01 at *Sac*II and *Kpn*I (pCS253). The 1.5 kb fragment was cloned into pPD117.01 at *Sac*II and *Not*I (pCS255). The 4.0 kb fragment was cloned into pPD117.01 at *Sac*II and *Not*I sites (pCS256).

dsRNAi

The templates used were the *yk1285a11* cDNA clone [containing sequences from predicted exons 1-7 and alternative exon 1 (Fig. 3)], and the *yk228h6* 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]. *yk1285a11* was digested by *Eco*RI and *Kpn*I; *yk228h6* was digested by *Sma*I and *Kpn*I. Then the digested DNA was extracted by phenol:chloroform once, precipitated by ethanol, dried in air, and dissolved in TE. 1μg of the cut DNA was used to synthesize RNA by Stratagene RNA Transcription Kit. After that, the reaction solution was treated by RNase-free DNaseI 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 10μl 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 pJKL547.1, 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 bacterially generated fusion proteins. For immunostaining, wild-type N2 and *wk55* mutant animals were fixed following the protocol of Hurd and Kemphues (Hurd and Kemphues, 2003). Rat anti-SMA-9 isoform II antiserum (serum CUMC-RT TB2) 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 (Collet 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 (*bxIs14*) (Barr and Sternberg, 1999) (L. Jia and S.W.E., unpublished). 5HT-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*(*lf*) 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 *sma-9(wk55)* mutants, which display a strong mutant phenotype.

sma-9 encodes a Shn-like protein

We mapped *sma-9* by positional cloning using standard genetic markers and SNPs. The mapping data placed *sma-9* on linkage group 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 *sma-9(wk55)* (*lf*) mutant background and assessed their ability to rescue body size and male tail defects. Cosmid T05A10 rescued *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, T05A10.1 (GenBank Accession Number Z68108) is predicted to encode a transcription factor homologous to *Drosophila* Shn, a Smad cofactor in the Dpp pathway. We hypothesized that T05A10.1 corresponds to *sma-9*. Sequencing of the corresponding ORF in *sma-9(wk55)* and *sma-9(qc3)* genomic DNA identified nonsense mutations in each (Fig. 2A): *wk55* converts Arg1163 to a stop codon and *qc3* converts Gln204 to a stop codon. The nonsense mutations in *wk55* and *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 *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 *sma-9*.

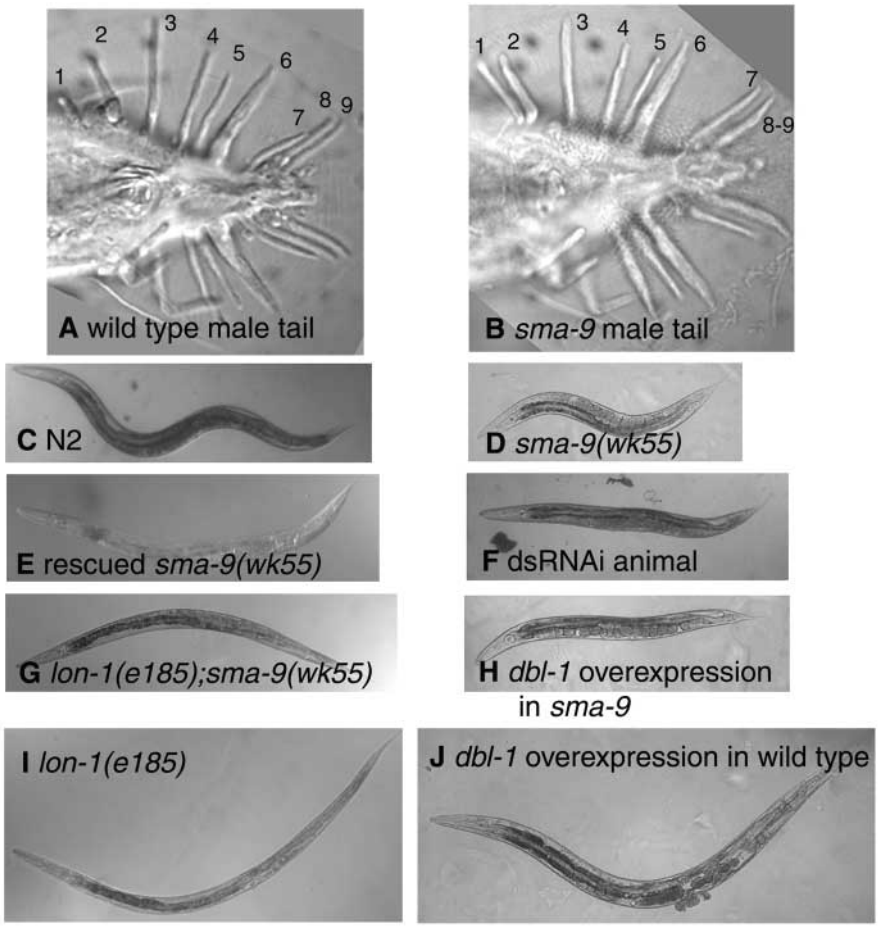

Fig. 1. *sma-9* mutants display Sma and Mab phenotypes similar to those of DBL-1 pathway mutants. (A,B) Male tail phenotype. Wild-type male tail (A) has nine bilateral pairs of sensory rays; *sma-9(wk55)* mutant (B) displays a ray 8-9 fusion. (C-F) Body size phenotype. *sma-9*(*lf*) mutant (D) is Sma compared with N2 (C). Cosmid T05A10 rescued *sma-9(wk55)* body size (E). (F) dsRNAi of *sma-9* 3′ exons in *him-5*. (G-J) Genetic interactions. *lon-1;sma-9* double mutant (G) is neither Lon nor Sma. *dbl-1* overexpression in wild-type background is Lon (J). However, the animal displays a Sma phenotype in *sma-9(wk55)* background (H). (I) *lon-1* single mutant. All animals in C-J are young adults photographed at the same magnification, with anterior oriented to the left.

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 C_2H_2 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 *sma-9* shares high sequence homology with a zinc finger transcription factor family that includes *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 αA-crystallin-binding protein 1 (αA-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 αA-CRYBP1 (Fig. 2B). Therefore, SMA-9, MBP1, αA-CRYBP1 and Shn may derive from a common ancestral gene, and this pair of zinc fingers may

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

Mutant	Percentage fusion of rays 4-5	Percentage fusion of rays 6-7	Percentage fusion of rays 8-9	*n* (sides)	Percentage crumpled spicules	*n* (tails)
Wild type	0	0	16	74	0	30
Single mutants						
sma-9(wk55)	0	0	49	105	46	50
sma-9(wk62)	0	0	35	102	66	49
sma-9(wk71)	0	0	24	104	32	52
sma-9(wk82)	0	0	34	103	43	51
sma-9(qc1)	0	0	67	30	32	20
sma-9(qc3)	0	0	29	100	7	37
sma-9(qc5)	0	0	34	59	65	20
sma-9(qc6)	0	0	31	80	53	53
sma-9(qc7)	0	0	51	112	55	80
sma-9(qc8)	0	0	43	86	63	47
sma-9(qc9)	0	0	33	93	58	34
sma-9(qc10)	0	0	47	51	88	16
sma-9(qc11)	0	0	26	105	51	88
sma-3(wk30)	26	64	21	121	100	30
Double mutants						
sma-3(wk30);sma-9(wk55)	11	40	16	81	ND	ND
sma-3(wk30);sma-9(wk62)	30	45	21	100	100	30
sma-9 rescue						
sma-9(wk55);qcEx49	0	1	23	79	9	42
dsRNAi						
exons 1-7	0	0	13	102	3	36
exons 21-25	0	0	26	70	4	46
exons 1-7 and 21-25	0	0	28	75	ND	ND

All strains contain *him-5(e1490)*.
qcEx49 contains cosmid T05A10 and plasmid pRF4.
ND, not determined.

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 αA-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

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, UUAUCCCUUUGUGUUUAAUU, 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 Sma 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 product 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.

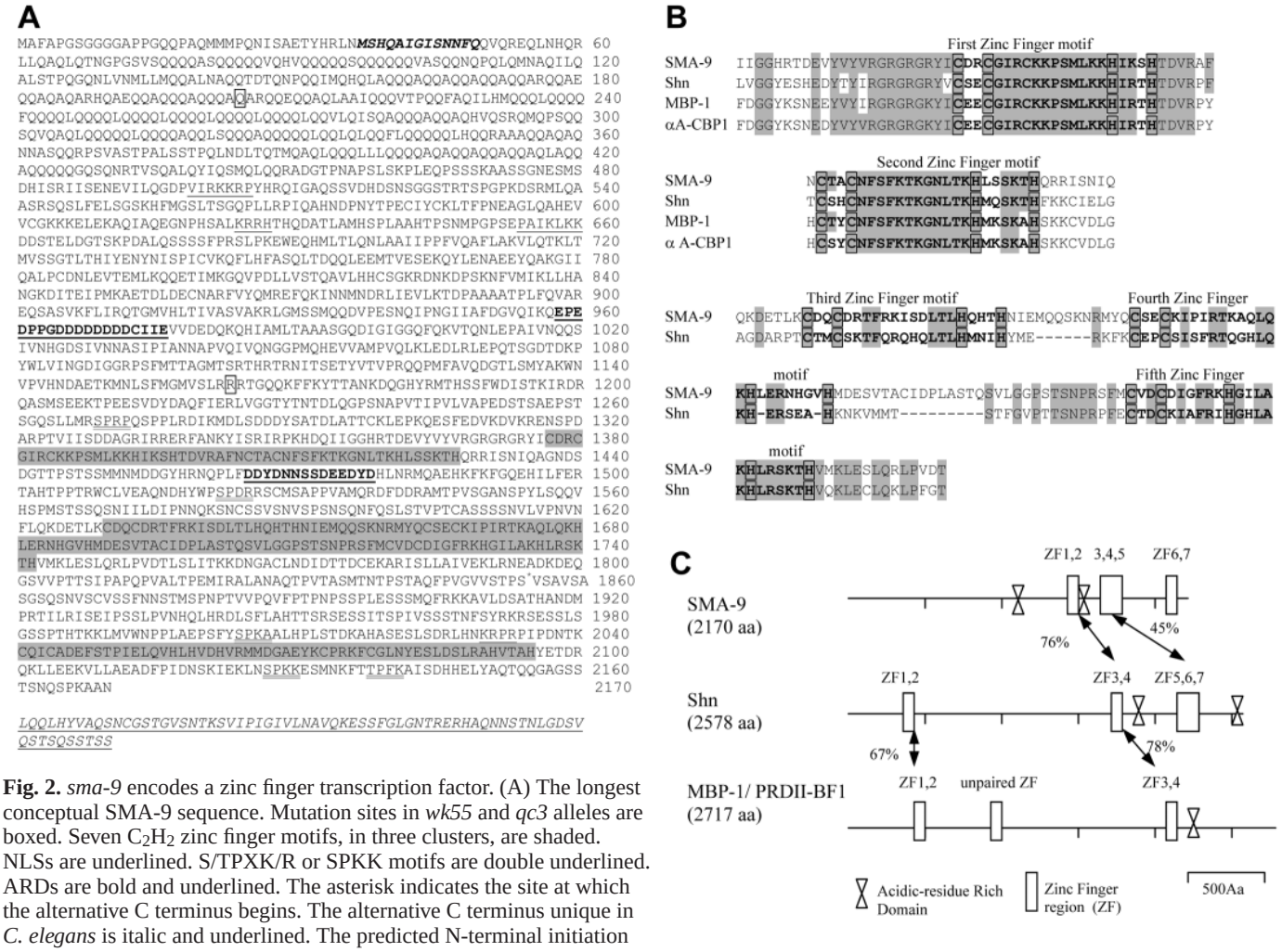

Fig. 2. *sma-9* encodes a zinc finger transcription factor. (A) The longest conceptual SMA-9 sequence. Mutation sites in *wk55* and *qc3* alleles are boxed. Seven C_2H_2 zinc finger motifs, in three clusters, are shaded. NLSs are underlined. S/TPXK/R or SPKK motifs are double underlined. ARDs are bold and underlined. The asterisk indicates the site at which the alternative C terminus begins. The alternative C terminus unique in *C. elegans* is italic and underlined. The predicted N-terminal initiation sequence (predicted exon 1) is italic and bold. (B) Comparison of the SMA-9, Shn, MBP1 and αA-CBP1 zinc finger regions. The Cys and His residues of zinc finger motifs are boxed. The identical residues are shaded. (C) Comparison of the SMA-9, Shn and MBP1 zinc finger domains. The SMA-9 first pair of zinc fingers is highly conserved between invertebrates and vertebrates. The SMA-9 triplet appears to have been either eliminated during vertebrate evolution or acquired in worm-fly lineage. The SMA-9 second pair is unique in *C. elegans*. The Shn first pair is conserved in vertebrate homologs but not in *C. elegans*.

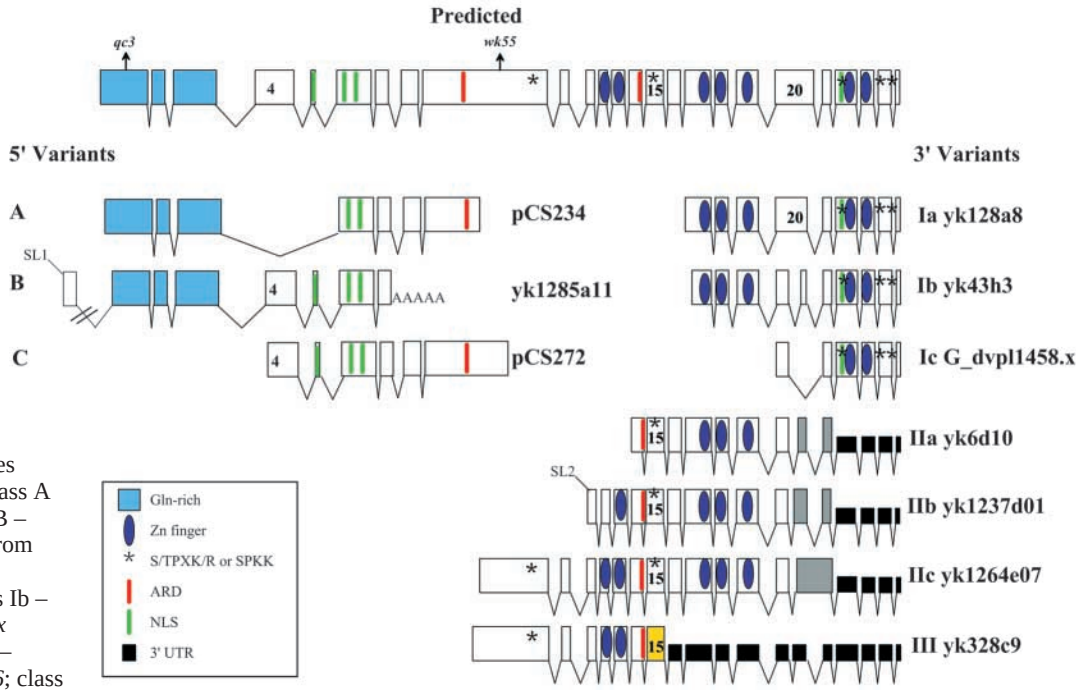

Fig. 3. Alternative splicing of *sma-9* transcripts. *sma-9* cDNA clones reveal a complex alternative splicing at both 5′ and 3′ ends. At least three N terminus and seven C terminus forms were found. Predicted exons 4, 15 and 20 are labeled. Approximate locations of the termination codons in *qc3* and *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 – *yk1285a11*; 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*, *yk864c1* and *yk1134e06*; class IIb – *yk1237d01*, which is SL2-spliced, *yk856b10*, *yk1057a6*, *yk1216e10*, *yk1103h10* and *yk127d10*, which is also missing the intron between exon 21 and 22 (not shown); class IIc – *yk1264e07*; and class III – *yk328c9*. Only the trans-spliced cDNA clones *yk1285a11* and *yk1237d01* appear to be full length.

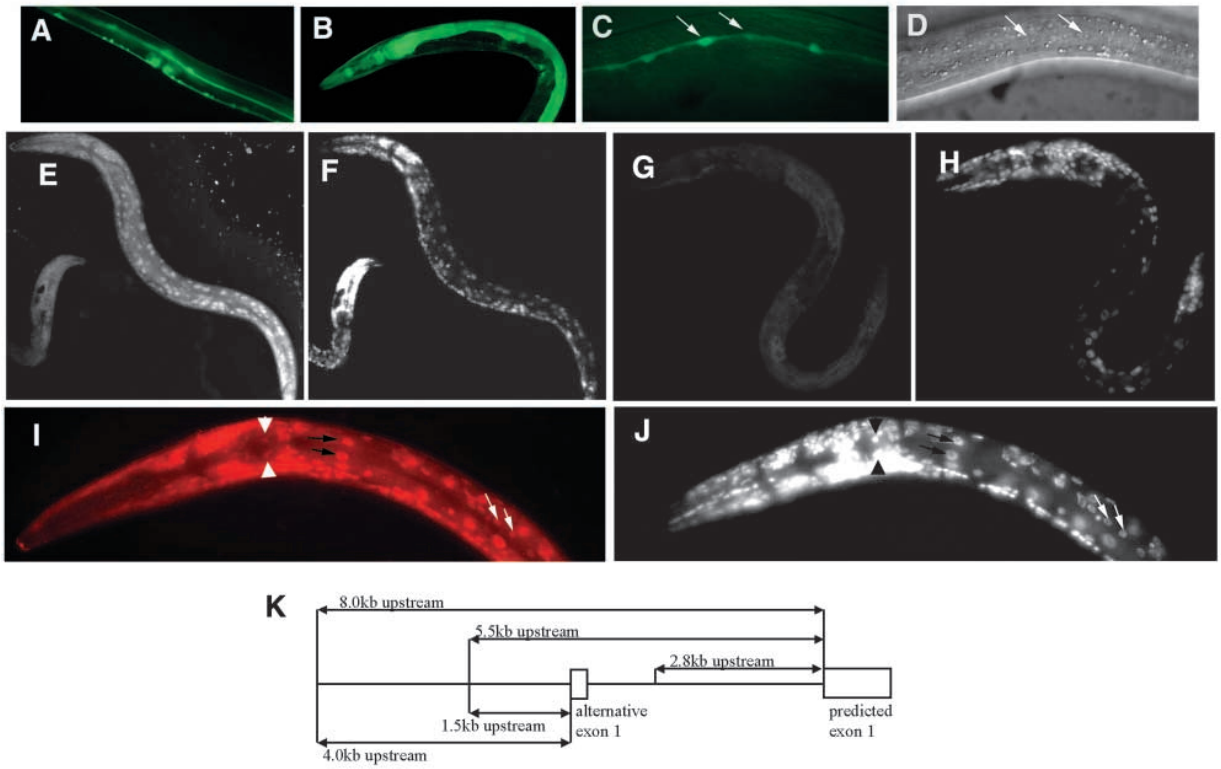

Fig. 4. *sma-9* is widely expressed. *sma-9* upstream sequences (K) were fused with a GFP reporter gene. (A-C) GFP expression in wild-type transgenic animals. The transgenic animals display a strong fluorescence in the VNC, pharynx and intestine, from stages L1 (not shown) to adult (A,B). L2 stage (C) 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 DBL-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.

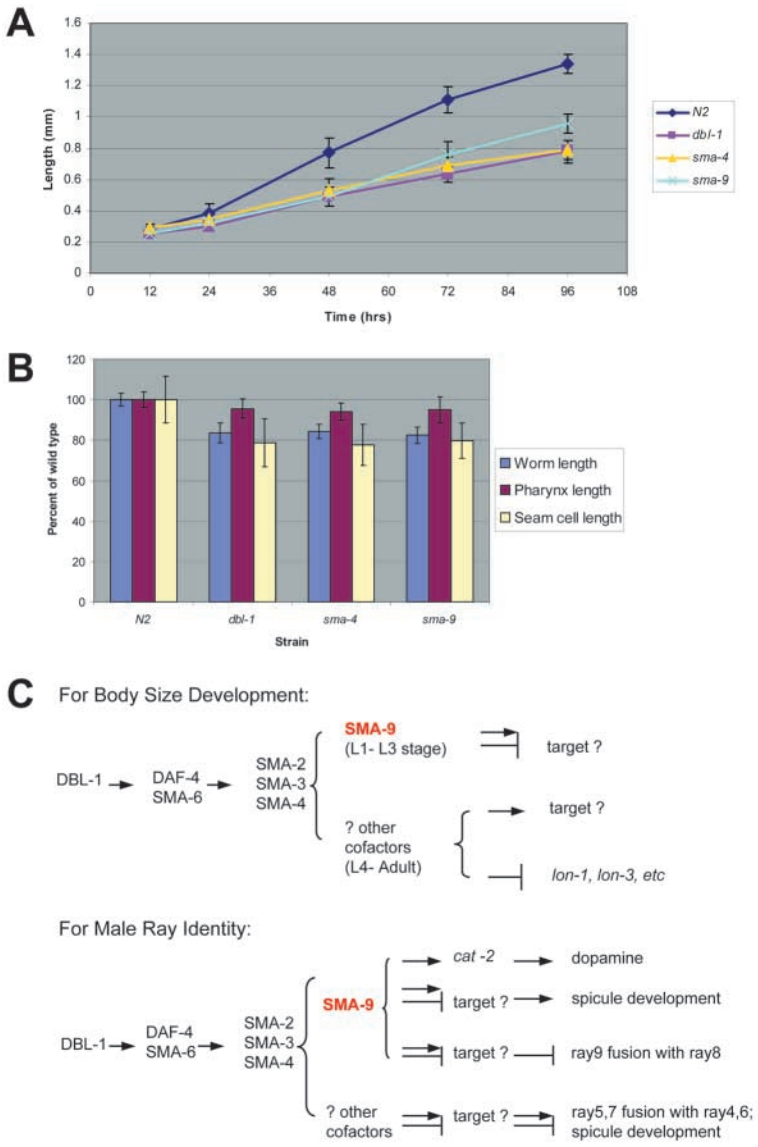

Fig. 5. *sma-9* functions in early stages to regulate body size. (A) Growth curve of *sma-9* (cross), *sma-4* (triangle), *dbl-1*(square) and N2 (diamond). In L1 stage larvae (12 hours), all animals are the same length. In L1 through L3 stages (48 hours), *sma-9*, *sma-4* and *dbl-1* show reduced a growth rate compared with N2. After L3, *sma-9* mutants grow rapidly, whereas *dbl-1* and *sma-4* mutants continue to display a reduced growth rate. Data for N2 and *dbl-1* is from Savage-Dunn et al. (Savage-Dunn et al., 2000). (B) At L3 stage, *sma-9* worm length, pharynx length and seam cell length is indistinguishable from that of *dbl-1* and *sma-4*. Data for N2 and *dbl-1* is from Wang et al. (Wang et al., 2002). (C) Model of *sma-9* function in the DBL-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 functions downstream of DBL-1 to regulate body size

The DBL-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 DBL-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 DBL-1 pathway mutants (Fig. 5A). In L2 and L3 stages, *sma-9* animals have the same size and growth rate as the DBL-1 pathway mutants have. Furthermore, the *sma-9*(*lf*) pharynx and seam cell lengths are indistinguishable from those of DBL-1 pathway mutants at the L3 stage (Fig. 5B). By contrast, another Sma mutant, *sma-1* (McKeown et al., 1998), which does not participate in this pathway, has a very different distribution of cell and organ sizes (Wang et al., 2002). However, after L3, *sma-9* mutant growth rate increases to a wild-type rate and finally gives the adult animals an intermediate body length (Fig. 5A). Therefore, although the DBL-1 pathway regulates body size throughout post-embryonic development,

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 frequencies 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.

Table 2. Percentage frequency of dopaminergic (DA) and serotonergic (5HT) fate expression in neurons of the male rays

Genotype	Neuron fate	Ray 1	2	3	4	5	6	7	8	9
Wild type	DA	–	–	–	–	100	–	100	–	100
sma-3 (e491)	DA	–	–	–	–	10	–	30	–	–
sma-9 (wk55)	DA	–	–	–	–	10*	–	10*	–	20*
sma-3; sma-9	DA					ND				
Wild type	5HT	100	–	100	–	–	–	–	–	100
sma-3 (e491)	5HT	100	–	100	–	100	–	–	–	–
sma-9 (wk55)	5HT	100	–	100	–	5*	–	5*	–	75
sma-3; sma-9	5HT	100	–	100	–	100	–	–	–	–

Percentage frequencies for each genotype correspond to (the number of times a ray was observed to express the neurotransmitter fate indicated/the total number of male sides scored, *n*) × 100%. 30 to 400 sides were scored/genotype.

Mutants defining all known *dbl-1* pathway genes were examined; representative examples are shown.

Expression of dopaminergic fate (in ray A-type neurons) was monitored using reporters for the dopamine biosynthesis gene *cat-2* (tyrosine hydroxylase) and/or by staining for dopamine directly using formaldehyde induced fluorescence (Lints and Emmons, 1999). Expression of serotonergic fate (in ray B-type neurons) was monitored using reporters for the serotonin biosynthesis gene *tph-1* (trytophan hydroxylase) and/or by staining with anti-serotonin antibodies (see Materials and methods) (Loer and Kenyon, 1993; Sze et al., 2000).

In *dbl-1* pathway and *sma-9* mutant backgrounds, percentage frequencies for ray 5, 7 or 9 also include fusions involving these rays.

Asterisk indicates weak expression of neurotransmitter markers.

ND, not determined.

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/PRDII-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)*

		Ray								
Genotype	Treatment	1	2	3	4	5	6	7	8	9
Wild type	no heat-shock	–	–	–	–	100	–	100	–	100
	+ heat-shock	–	–	40	60	100	20	100	20	100
sma-9 (wk55)	no heat-shock	–	–	–	–	10*	–	10*	–	15*
	+ heat-shock	–	–	–	–	10*	5*	10*	–	18*

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.

mutants, their genetic interactions, the nuclear localization of endogenous SMA-9 and the overlapping expression with DBL-1 pathway genes, we propose that SMA-9 functions as a Smad transcriptional cofactor in the DBL-1 pathway.

The similarity of specific aspects of the *sma-9* and DBL-1 pathway mutant Sma and Mab phenotypes provides strong support that *sma-9* functions in the DBL-1 BMP-related pathway. Like DBL-1 pathway mutants, *sma-9*(*lf*) mutants grow slowly during post-embryonic development, and have reduced body length and seam cell size (Fig. 5). From stages L1 to L3, the *sma-9* mutant body size is indistinguishable from that of DBL-1 pathway mutants. In the male rays both *sma-9* and DBL-1 signaling induce DA fate expression in A-type neurons of rays 5, 7 and 9; in the B-type neurons they suppress inappropriate expression of serotonergic fate in ray 5 and induce this fate in ray 9 (Table 2). Furthermore, *sma-9(lf)* can suppress the effects of *dbl-1* overexpression in both body size and male tail phenotypes (Fig. 1, Table 3). These experiments demonstrate that *sma-9* acts genetically downstream in the DBL-1 pathway.

sma-9 functions in a temporally and spatially specific manner

sma-9(*lf*) mutants display a stage-specific body size phenotype. *sma-9* mutant growth rate indicates that the wild-type gene product is required for body growth from the L1 to L3 stage, but is dispensable at later stages (Fig. 5). As DBL-1 pathway mutants continue to grow slowly throughout development, this suggests that after L3 some other cofactors might replace *sma-9*. In contrast to *lon-1* and *lon-3*, which function at late larval stages (Maduzia et al., 2002; Morita et al., 2002; Nystrom et al., 2002; Suzuki et al., 2002), *sma-9* is the first DBL-1 pathway component contributing only to early larval stage body size development. The large hypodermal syncytium hyp7 has been shown to be a crucial tissue for body size regulation (Yoshida et al., 2001; Wang et al., 2002; Maduzia et al., 2002; Morita et al., 2002), and is a site of expression for *sma-6* (Krishna et al., 1999), *daf-4* (Gunther et al., 2000) and *sma-3* (Savage-Dunn et al., 2000). SMA-9 is similarly detected in nuclei of hyp7. Furthermore, the expression of *sma-9* promoter-driven GFP is detected in the seam cells from stages L1 to L3, coincident with the requirement for *sma-9* in body size regulation. The stage-specific *sma-9* expression in the lateral seam suggests that the seam cells, together with hyp7, play an important role in body size regulation by the DBL-1 pathway.

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*(*lf*) 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*(*lf*) 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 Maduzia and Richard Padgett for the BW1940 *ctIs40* 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 Daryl Hurd 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 RPG 98-230-01-DDC from the American Cancer Society and a PSC-CUNY research award to C.S.-D.; by NIH R01 GM066953 to J. Liu; and by NIH R01 NS38906 to S.W.E. M.L.F. was partially supported by NIH T32 GM07617; R.L. is a NARSAD Young Investigator; and S.W.E. is the Siegfried Ullmann Professor of Molecular Genetics. Some *C. elegans* mutant strains were obtained from the Caenorhabditis Genetics Center, which is supported by the NIH National Center for Research Resources (NCRR). This work was carried out in partial fulfillment of the requirements for a PhD degree from the Graduate Center of the City University of New York (J. Liang).

References

Arora, K., Dai, H., Kazuko, S. G., Jamal, J., O'Connor, M. B., Letsou, A. and Warrior, R. (1995). The Drosophila schnurri gene acts in the Dpp/TGF beta signaling pathway and encodes a transcription factor homologous to the human MBP family. *Cell* **81**, 781-790.

Barr, M. M. and Sternberg, P. W. (1999). A polycystic kidney-disease gene homologue required for male mating behaviour in C. elegans. *Nature* **401**, 386-389.

Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J., Thierry-Mieg, D., Chiu, W. L., Duke, K., Kiraly, M. and Kim, S. K. (2002). A global analysis of Caenorhabditis elegans operons. *Nature* **417**, 851-854.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94.

Campbell, G. and Tomlinson, A. (1999). Transducing the Dpp morphogen gradient in the wing of Drosophila: regulation of Dpp targets by brinker. *Cell* **96**, 553-562.

Chen, X., Rubock, M. J. and Whitman, M. (1996). A transcriptional partner for Mad proteins in TGF-β signalling. *Nature* **389**, 691-696.

Collet, J., Spike, C. A., Lundquist, E. A., Shaw, J. E. and Herman, R. K. (1998). Analysis of *osm-6*, a gene that affects sensory cilium structure and sensory neuron function in *Caenorhabditis elegans*. *Genetics* **148**, 187-200.

Dai, H., Hogan, C., Gopalakrishnan, B., Torres-Vazquez, J., Nguyen, M., Park, S., Raftery, L. A., Warrior, R. and Arora, K. (2000). The zinc finger protein schnurri acts as a Smad partner in mediating the transcriptional response to decapentaplegic. *Dev. Biol*. **227**, 373-387.

Estevez, M., Attisano, L., Wrana, J. L., Albert, P. S., Massague, J. and Riddle, D. L. (1993). The daf-4 gene encodes a bone morphogenetic protein receptor controlling C. elegans dauer larva development. *Nature* **365**, 644-649.

Fan, C. M. and Maniatis, T. (1990). A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. *Genes Dev*. **4**, 29-42.

Gascoigne, N. R. (2001). Positive selection in a Schnurri. *Nat. Immunol*. **2**, 989-991.

Grieder, N. C., Nellen, D., Burke, R., Basler, K. and Affolter, M. (1995). Schnurri is required for Drosophila Dpp signaling and encodes a zinc finger protein similar to the mammalian transcription factor PRDII-BF1. *Cell* **81**, 791-800.

Gunther, C. V., Georgi, L. L. and Riddle, D. L. (2000). A Caenorhabditis elegans type I TGF beta receptor can function in the absence of type II kinase to promote larval development. *Development* **127**, 3337-3347.

Hata, A., Seoane, J., Lagna, G., Montalvo, E., Hemmati-Brivanlou, A. and Massague, J. (2000). OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-Smad and Olf signaling pathways. *Cell* **100**, 229-240.

Hill, C. S., Packman, L. C. and Thomas, J. O. (1990). Phosphorylation at clustered -Ser-Pro-X-Lys/Arg- motifs in sperm-specific histones H1 and H2B. *EMBO J*. **9**, 805-813.

Hogan, B. L. (1996). Bone morphogenetic proteins in development. *Curr. Opin. Genet. Dev*. **6**, 432-438.

Huang, T., Kuersten, S., Deshpande, A. M., Spieth, J., MacMorris, M. and Blumenthal, T. (2001). Intercistronic region required for polycistronic pre-mRNA processing in *Caenorhabditis elegans*. *Mol. Cell. Biol*. **21**, 1111-1120.

Hurd, D. D. and Kemphues, K. J. (2003). PAR-1 is required for morphogenesis of the Caenorhabditis elegans vulva. *Dev. Biol*. **253**, 54-65.

Jazwinska, A., Kirov, N., Wieschaus, E., Roth, S. and Rushlow, C. (1999). The Drosophila gene brinker reveals a novel mechanism of Dpp target gene regulation. *Cell* **96**, 563-573.

Krishna, S., Maduzia, L. L. and Padgett, R. W. (1999). Specificity of TGFbeta signaling is conferred by distinct type I receptors and their associated SMAD proteins in Caenorhabditis elegans. *Development* **126**, 251-260.

Lallemand, C., Palmieri, M., Blanchard, B., Meritet, J. F. and Tovey, M. G. (2002). GAAP-1: a transcriptional activator of p53 and IRF-1 possesses pro-apoptotic activity. *EMBO Rep*. **3**, 153-158.

Lints, R. and Emmons, S. W. (1999). Patterning of dopaminergic neurotransmitter identity among Caenorhabditis elegans ray sensory neurons by a TGFbeta family signaling pathway and a Hox gene. *Development* **126**, 5819-5831.

Lints, R. and Emmons, S. W. (2002). Regulation of sex-specific differentiation and mating behavior in C. elegans by a new member of the DM domain transcription factor family. *Genes Dev*. **16**, 2390-2402.

Liu, Z., Kirch, S. and Ambros, V. (1995). The *C. elegans* heterodronic gene pathway controls stage-specific transcription of collagen genes. *Development* **121**, 2471-2478.

Loer, C. M. and Kenyon, C. J. (1993). Serotonin-deficient mutants and male mating behavior in the nematode *Caenorhabditis elegans*. *J. Neurosci*. **13**, 5407-5417.

Maduzia, L. L., Gumienny, T. L., Zimmerman, C. M., Wang, H., Shetgiri, P., Krishna, S., Roberts, A. F. and Padgett, R. W. (2002). lon-1 regulates Caenorhabditis elegans body size downstream of the dbl-1 TGF beta signaling pathway. *Dev. Biol*. **246**, 418-428.

Marty, T., Muller, B., Basler, K. and Affolter, M. (2000). Schnurri mediates Dpp-dependent repression of brinker transcription. *Nat. Cell Biol*. **2**, 745-749.

Massague, J. (1998). TGF-beta signal transduction. *Annu. Rev. Biochem*. **67**, 753-791.

McKeown, C., Praitis, V. and Austin, J. (1998). sma-1 encodes a betaH-spectrin homolog required for Caenorhabditis elegans morphogenesis. *Development* **125**, 2087-2098.

Mello, C. C., Kramer, J. M., Stinchcomb, D. T. and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J*. **10**, 3959-3970.

Morita, K., Chow, K. L. and Ueno, N. (1999). Regulation of body length and male tail ray pattern formation of Caenorhabditis elegans by a member of TGF-beta family. *Development* **126**, 1337-1347.

Morita, K., Flemming, A. J., Sugihara, Y., Mochii, M., Suzuki, Y., Yoshida, S., Wood, W. B., Kohara, Y., Leroi, A. M. and Ueno, N. (2002). A Caenorhabditis elegans TGF-beta, DBL-1, controls the expression of LON-1, a PR-related protein, that regulates polyploidization and body length. *EMBO J*. **21**, 1063-1073.

Muller, B., Hartmann, B., Pyrowolakis, G., Affolter, M. and Basler, K. (2003). Conversion of an extracellular Dpp/BMP morphogen gradient into an inverse transcriptional gradient. *Cell*. **113**, 221-233.

Nakamura, T., Donovan, D. M., Hamada, K., Sax, C. M., Norman, B., Flanagan, J. R., Ozato, K., Westphal, H. and Piatigorsky, J. (1990). Regulation of the mouse alpha A-crystallin gene: isolation of a cDNA

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

Nakayama, T., Cui, Y. and Christian, J. L. (2000). Regulation of BMP/Dpp signaling during embryonic development. *Cell. Mol. Life Sci*. **57**, 943-956.

Nystrom, J., Shen, Z. Z., Aili, M., Flemming, A. J., Leroi, A. and Tuck, S. (2002). Increased or decreased levels of Caenorhabditis elegans lon-3, a gene encoding a collagen, cause reciprocal changes in body length. *Genetics* **161**, 83-97.

Oukka, M., Kim, S. T., Lugo, G., Sun, J., Wu, L. C. and Glimcher, L. H. (2002). A mammalian homolog of Drosophila schnurri, KRC, regulates TNF receptor-driven responses and interacts with TRAF2. *Mol. Cell*. **9**, 121-131.

Patterson, G. I. and Padgett, R. W. (2000). TGF beta-related pathways. Roles in Caenorhabditis elegans development. *Trends Genet*. **16**, 27-33.

Randall, R. A., Germain, S., Inman, G. J., Bates, P. A. and Hill, C. S. (2002). Different Smad2 partners bind a common hydrophobic pocket in Smad2 via a defined proline-rich motif. *EMBO J*. **21**, 145-156.

Savage, C., Das, P., Finelli, A. L., Townsend, S. R., Sun, C. Y., Baird, S. E. and Padgett, R. W. (1996). Caenorhabditis elegans genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor beta pathway components. *Proc. Natl. Acad. Sci. USA* **93**, 790-794.

Savage-Dunn, C. (2001). Targets of TGF beta-related signaling in Caenorhabditis elegans. *Cytokine Growth Factor Rev*. **12**, 305-312.

Savage-Dunn, C., Tokarz, R., Wang, H., Cohen, S., Giannikas, C. and Padgett, R. W. (2000). SMA-3 Smad has specific and critical functions in DBL-1/SMA-6 TGFbeta-related signaling. *Dev. Biol*. **223**, 70-76.

Savage-Dunn, C., Maduzia, L. L., Zimmerman, C. M., Roberts, A. F., Cohen, S., Tokarz, R. and Padgett, R. W. (2003). A genetic screen for small body size mutants in *C. elegans* reveals many TGFβ pathway components. *Genesis* **35**, 239-247.

Seeler, J. S., Muchardt, C., Suessle, A. and Gaynor, R. B. (1994). Transcription factor PRDII-BF1 activates human immunodeficiency virus type 1 gene expression. *J. Virol*. **68**, 1002-1009.

Shi, Y., Wang, Y. F., Jayaraman, L., Yang, H., Massague, J. and Pavletich, N. P. (1998). Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling. *Cell* **94**, 585-594.

Staehling-Hampton, K., Laughon, A. S. and Hoffmann, F. M. (1995). A Drosophila protein related to the human zinc finger transcription factor PRDII/MBPI/HIV-EP1 is required for dpp signaling. *Development* **121**, 3393-3403.

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Dev. Biol*. **56**, 110-156.

Suzuki, Y., Yandell, M. D., Roy, P. J., Krishna, S., Savage-Dunn, C., Ross, R. M., Padgett, R. W. and Wood, W. B. (1999). A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in Caenorhabditis elegans. *Development* **126**, 241-250.

Suzuki, Y., Morris, G. A., Han, M. and Wood, W. B. (2002). A cuticle collagen encoded by the lon-3 gene may be a target of TGF-beta signaling in determining caenorhabditis elegans body shape. *Genetics* **162**, 1631-1639.

Sze, J. Y., Victor, M., Loer, C., Shi, Y. and Ruvkun, G. (2000). Food and metabolic signaling defects in a *Caenorhabditis elegans* serotonin-synthesis mutant. *Nature* **403**, 560-564.

Takagi, T., Harada, J. and Ishii, S. (2001). Murine Schnurri-2 is required for positive selection of thymocytes. *Nat. Immunol*. **2**, 1048-1053.

Torres-Vazquez, J., Park, S., Warrior, R. and Arora, K. (2001). The transcription factor Schnurri plays a dual role in mediating Dpp signaling during embryogenesis. *Development* **128**, 1657-1670.

van't Veer, L. J., Lutz, P. M., Isselbacher, K. J. and Bernards, R. (1992). Structure and expression of major histocompatibility complex-binding protein 2, a 275-kDa zinc finger protein that binds to an enhancer of major histocompatibility complex class I genes. *Proc. Natl. Acad. Sci. USA* **89**, 8971-8975.

Walhout, A. J., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N. and Vidal, M. (2000). Protein interaction mapping in C. elegans using proteins involved in vulval development. *Science* **287**, 116-122.

Wang, J., Tokarz, R. and Savage-Dunn, C. (2002). The expression of TGFbeta signal transducers in the hypodermis regulates body size in C. elegans. *Development* **129**, 4989-4998.

Yoshida, S., Morita, K., Mochii, M. and Ueno, N. (2001). Hypodermal expression of Caenorhabditis elegans TGF-beta type I receptor SMA-6 is essential for the growth and maintenance of body length. *Dev. Biol*. **240**, 32-45.