SMA-9 IN TGF-BETA SIGNALING PATHWAY OF *CAENORHABDITIS ELEGANS*

by

JUN LIANG

A dissertation submitted to the Graduate Faculty in Biochemistry
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, The City University of New York

2005

UMI Number: 3187381

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the

Graduate Faculty in Biochemistry in satisfaction of the

dissertation requirement for the degree of Doctor of Philosophy

_____	_____
Date	Chair of Examining Committee
	Dr. Cathy Savage-Dunn
_____	_____
Date	Executive Officer
	Dr. Lesley Davenport
	Corinne A. Michels
	Timothy W. Short
	Thomas Haines
	Christine A. Rushlow
	Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

SMA-9 IN TGF-β SIGNALING PATHWAY OF *CAENORHABDITIS ELEGANS*

by

Jun Liang

Adviser: Professor Cathy Savage-Dunn

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, *sma-9* loss of function phenotype, and the expression pattern support that *sma-9* acts as a downstream component and is required in the DBL-1 signaling pathway, providing the first evidence for 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.

Our genetic studies using artificial SMA-9 constructs give the first evidence that a transcriptional repressor can substitute for functional Shn protein *in vivo.* We showed that the *sma-9* N-terminal Gln-rich region contributes to this transcriptional repressor activity by an *in vitro* transcriptional activity assay. Furthermore among the total genes regulated by *sma-9* and *dbl-1* from our microarray analysis, there are more repressed genes than

activated genes. Taken together, our results suggest that SMA-9 functions as a transcriptional repressor in DBL-1 pathway regulated body size development and male tail morphogenesis.

ACKNOWLEDGEMENTS

I greatly thank Dr. Cathy Savage-Dunn for her continuous support, encouragement and genuine advice in both my research project and career. During these five years, she gave me the freedom of trying new things although some did not turn out as planned, and encouraged my own interests and idea. Talking to Cathy is always joyful since I can always learn something. She showed me an excellent scientist and brilliant mentor. My committee members have been extremely supportive for reaching my goal throughout my Ph.D. and my career, in particular Dr. Michels and Dr. Haines.

The biochemistry program at Graduate center - CUNY provided me a great scientific training. I learned a lot from the courses, lectures, and seminars. This program is the best that I ever had, although passing the first level test was painful but fruitful. Professors in the biochemistry program showed me excellent models of what a scientist should be, in particular Dr. Horst Schulz and Dr. Lesley Davenport.

I owe many thanks to every person in Dr. Savage-Dunn lab for their help. In particularly, I greatly thank Ling Yu who taught me a lot of worm techniques. She has always helped me for this project. We shared happiness, as well as frustration. I thank Marie McGovern for helpful discussions. Jianjun Wang has been a great working partner; his passion in science encourages me and makes me fulfill the project.

The members in Dr. Michels and Dr. Short's lab are always very generous. I learned many tips from them for my yeast experiments; we shared ideas of molecular biology techniques. I thank Bing Zhang, Xing Wang, Fulai Ran, Caihong Qiu, Fang Qi, and all the other people in their labs.

Yuxing Dong and her family are always very kind to me. I remember the first but toughest year when I came to the New York City, they gave me encouragement to continue my study. From them, I see the future and gain the strength. Also, I am grateful to Dr. William J. Rice for his help with the manuscript.

Last but not least, I have to thank my wonderful family: my parents and my brother for all their unconditional support, love, and encouragement. They always worry about me being alone in this country, but they showed their best support. It is they who make me accomplish this much and make my dream come true.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

I. *C. elegans* is an excellent model organism for genetic study

Nematodes parasitize almost all kinds of plant and animal organisms. They live in all terrestrial and marine environments and feed on everything that gets into their mouths. *Caenorhabditis elegans* is a kind of soil nematode. *C. elegans* are free-living animals. In the laboratory, *C. elegans* is able to grow on petri dishes with *Escherichia coli* as a food source. The whole animal length at adulthood is about 1.5 mm.

C. elegans was first introduced for genetic studies by Sydney Brenner in 1965 in order to study animal behavior and development. They characterized the whole cell lineage and established a practical system for storage in the laboratory (Brenner, 1974). Their pioneering and detailed work made the animal a more and more popular model organism in life science. To date, worm research covers almost every corner of biology, for example, chemosensation, sex determination, signal transduction, neuron science, muscle contraction, and so on. There are several reasons that make *C. elegans* an excellent model organism for research.

First, it has a rapid life cycle: 3 days per generation at 20°C. A single hermaphrodite gives rise to 300-350 progeny, which provides a great advantage for genetic studies.

Second, a transparent body during its whole lifetime makes it easy to examine animal phenotypes under the microscope. As a result, it is not surprising that the complete wild-type *C. elegans* cell lineage has been determined from fertilized egg to adult.

Third, the cell number and cell position are nearly constant between individuals. This provides a very strong advantage for studying development and genetics compared with other model organisms.

Fourth, *C. elegans* is a simple animal. Each adult animal has less than 1,000 somatic cells, including 302 neuron cells in hermaphrodite or 381 in male. The animal has a small genome: about 97 million base pair. Approximately 19,099 genes are predicted and only about 10% of them have been studied genetically or biochemically (The C.elegans Sequencing Consortium, 1998). There are still a large number of genes that require to be characterized. This provides a big challenging space for researchers to explore.

Fifth, the coexistence of self fertilization and cross fertilization between hermaphrodite and male makes it convenient for genetic study. *C. elegans* exists mostly as a hermaphrodite that is able to produce both oocyte and sperm, which leads to self-fertilization. At the same time, males can cross with hermaphrodites. Under this situation, the oocyte prefers sperm from the male so that the diversity of genotypes increases. In the laboratory, males are often used to cross with hermaphrodites in order to create double or triple mutants for phenotype characterization.

Finally, another key feature is that *C. elegans* is good for transformation. Plasmid or cosmid DNA can be microinjected into the hermaphrodite gonad and taken by the germline cells. These DNAs act as an extra chromosome in the genome and can be passed through generations. Also γ-irradiation can be used to integrate these DNAs into the worm genome so that the genes can be stably expressed. Moreover, gene knockout techniques have been applied to the animal.

1. Life Cycle of *C. elegans*

C. elegans has six normal developmental stages: egg; the four larval stages L1, L2, L3, and L4; and adult. The larval stages can be distinguished by their specific cuticle secreted by the underlying hypodermis. As the old cuticle is shed new one is created. At 25°C, embryogenesis requires 14 hours; post-embryonic development from L1 to adult needs 36 hours. The rapid life cycle and large number of progeny produced help them survive in the earth by overgrowing their competitors.

In addition, the dauer stage is a special developmental stage with specialized survival strategies. Under tough living conditions, for example insufficient food or hot temperature etc, L2 animals enter a non-developmental stage - dauer stage – instead of L3. The dauer changes energy metabolism, accumulates fat, alters its behavior strikingly. At the same time, a specific cuticle replaces the L2 cuticle. Dauers are dark slim animals that either move very fast to seek a better living environment or stay in the same place waiting for environmental improvement. They can be recovered directly to L4 under appropriate conditions.

2. *C. elegans* Anatomy Structure

The animal body is composed of two concentric tubes (Fig A). The outer tube includes: body wall, gonad, part of nervous system, excretory, and secretory system. The inner tube includes: pharynx, the remaining nervous system, intestine, rectum, and anus. Most neurons are around the pharynx, along the ventral midline, and in the tail.

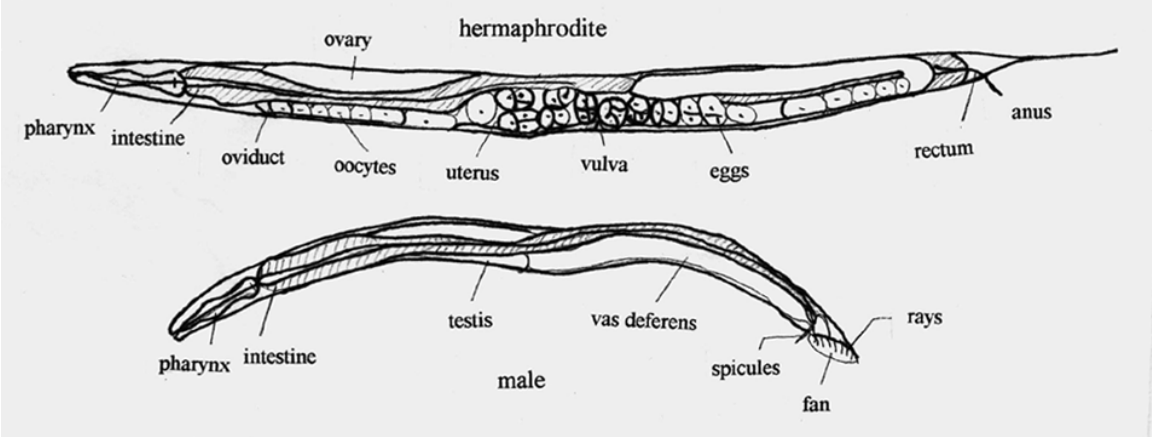


Fig A. Schematic structure of the *C. elegans* adult hermaphrodite and male (modified from Sulston and Horvitz, 1977)

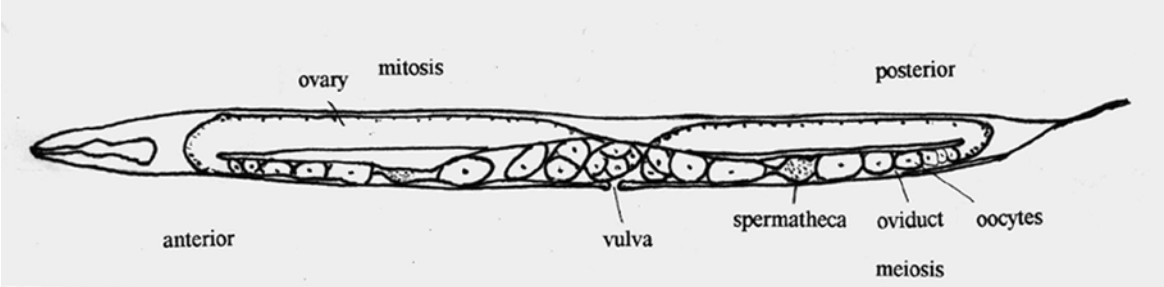


Fig B. Schematic structure of hermaphrodite gonad (modified from Hansen et al., 2004a; Miskowski et al., 2001)

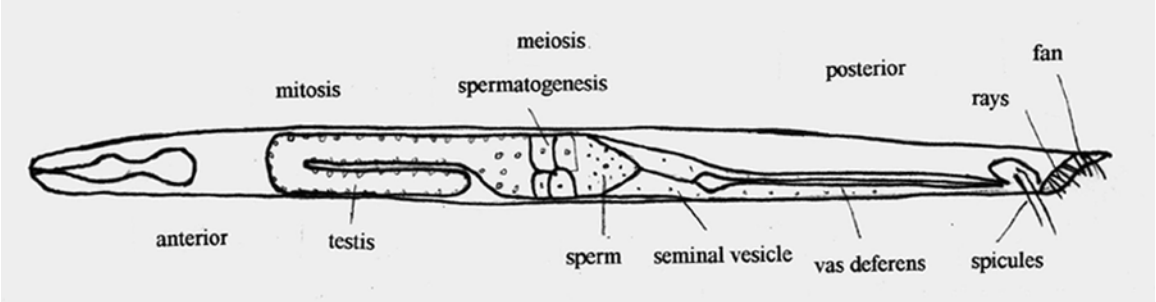


Fig C. Schematic structure of male gonad (Modified from Miskowski et al., 2001; Kimble and Hirsh, 1979)

The hermaphrodite gonad is made up of anterior and posterior arms, which is reflexed with an ovary (Fig B). In the ovary, the germ cells perform mitosis. Then meiosis occurs at the oviduct in order to give rise to oocytes. The sperm is made and stored in the spermatheca where each oocyte is fertilized when it passes through. The fertilized egg continues to develop to the early embryo stage in the uterus until it is time to be laid out (Hansen et al., 2004a; Lee and Schedl, 2001).

The male gonad has a single reflex (Fig C). The anterior arm is the testis where the germ cells perform mitosis and meiosis in order to produce mature sperm. The posterior arm includes: vas deferens; seminal vesicle; and cloaca.

The copulatory bursa in the male tail includes two spicules, a fan, and 18 sensory rays. The fan is an extended acellular structure of cuticle matrix. The 18 rays split into two identical groups that are one-axis symmetrical along the ventral midline. Each group contains 9 different rays: numbered 1 through 9 from anterior to posterior (Liang et al., 2003), which can be distinguished by their specific location and shape. Rays 1, 5, and 7 localize at the dorsal side of the fan; rays 3, 6, and 9 at the edge; rays 2, 4, and 8 at the ventral side. The function of the sensory ray is to sense contact with the hermaphrodite and regulate male turning during the mating. Each group of rays is independent even if they belong to the same individual. Within each ray there are two neurons, RNA and RNB, which express different neurotransmitters, and one support cell surrounded by a hypodermal sheath. For example, dopamine is expressed only in the ray5 A, ray7 A, and ray9 A neurons (Lints and Emmons, 1999). Two long sharp spicules locate at the male tail that insert into the hermaphrodite vulva, which locates the hermaphrodite vulva

precisely (Garcia et al., 2001). Since the male mating function depends on sensory rays and spicules, their misshape results in dysfunction of mating.

3. Hypodermal Cells

The animal body is surrounded by hypodermal cells and their secreted cuticle. The hypodermal cells are classified into four groups: hyp7, seam cells, hypodermal cells of the head and tail, and interfacial hypodermal cells. Because of cell fusion during development, many hypodermal cells in *C. elegans* are multinucleate.

Hyp7 is a single multinucleate cell or syncytium that covers the major part of the animal body (Fig D). At hatching, hyp7 contains 23 diploid nuclei and forms the dorsal hypodermis. Then, daughter cells from 12 ventral P blast cells and lateral seam cells undergo endoreplication and fuse into hyp7. In the adult, the nuclei are tetraploid. During the entire post-embryonic development, a total of 110 cells fuse together and form hyp7.

At hatching, ten seam cells: H0-H2, V1-V6, and T, form a row along each lateral side. All these cells apart of the most anterior one, are blast cells. During larval development, these blast cells divide into two daughter cells, one remains as a seam cell, while the other joins hyp7 to form body syncytium. In adults, two seam syncytia are made up by seam cell fusion; the structure is separate from main body syncytium hyp7.

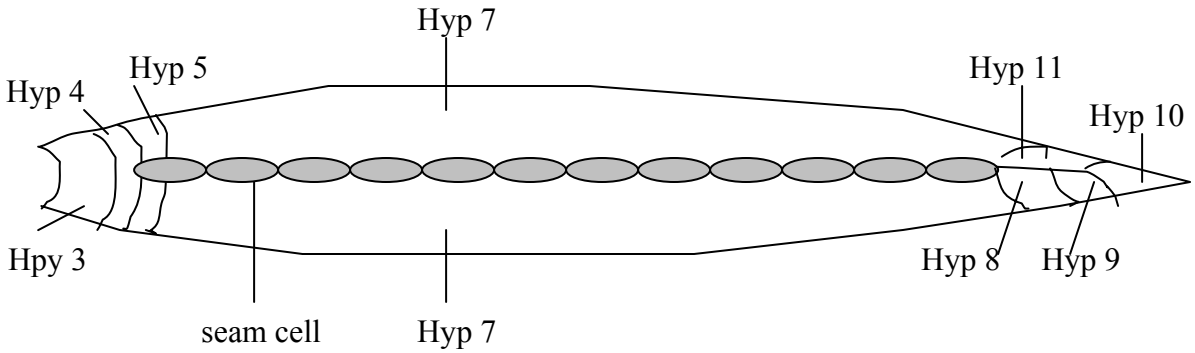


Fig D. Schematic structure of *C. elegans* hypodermal cells in hermaphrodite

The male seam cells are blast cells during larval development as they are in hermaphrodite. The anterior cells generate cuticles. However, the three most posterior seam cells: V5, V6, and T perform extra divisions that give rise to nine pairs of male ray cells (Rn) in the L3 stage. These Rn cells generate two sensory neurons RNA and RNB and one structural cell (Lints and Emmons, 1999; Lambie, 2002).

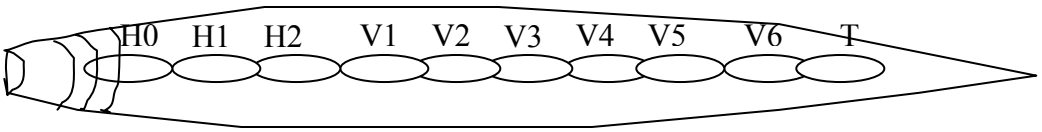


Fig E. lateral epidermal seam cells in L1 male

II. TGF-β signal transduction

The transforming growth factor-β (TGF-β) superfamily is a large group of secreted growth factors that regulates various cell responses such as cell arrest, cell proliferation, apoptosis, cell differentiation etc. It exists both in vertebrates and invertebrates, including TGF-β_1, TGF-β_2, TGF-β_3, Activins, bone morphogenetic proteins (BMPs) etc (Massague, 1998; ten Dijke and Hill, 2004). TGF-β ligands play a critical role in development in many organisms, such as limb formation in mouse (Capdevila and Izpisua Belmonte, 2001), mesoderm induction in frog (Grimm OH and Gurdon JB 2002), appendage development in *Drosophila* (Tabata and Takei, 2004), body size and male tail morphogenesis in *C. elegans* etc (Savage et al., 1996).

TGF-β signaling starts from association of ligand with two transmembrane ser/thr kinase receptors – type I receptor and type II receptor, and then the ligand-receptor complex phosphorylates R-Smad protein (Receptor- regulated Smads) (Fig F). After that, activated R-Smads recruit Co-Smad (common-mediator Smads) and form a Smad complex that is translocated into the nucleus in order to activate or repress gene transcription depending on whether it associates with transcriptional activator or repressor. Compared with other growth factor pathways, TGF-β signaling is comparatively simple, therefore it requires more regulatory components in order to output specific cell responses.

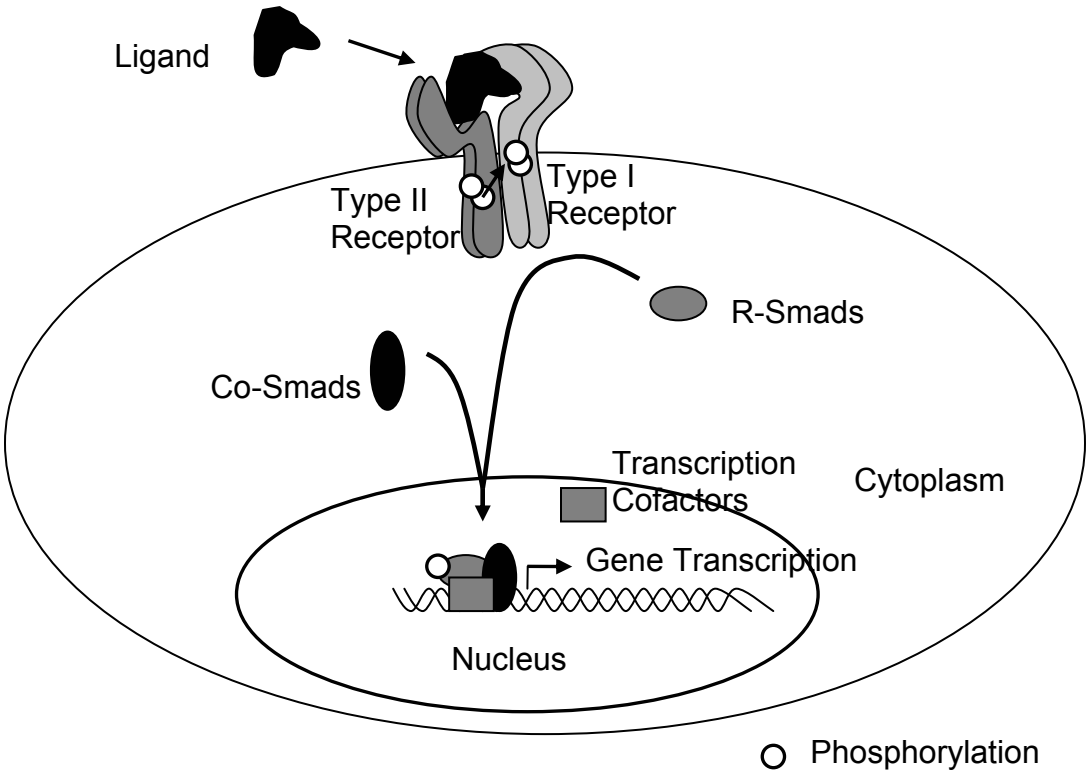


Fig F. TGF-β Signaling Pathway

1. TGF-β - the Ligand

Phylogenetic analysis categorizes the TGF-β superfamily into two large subfamilies, TGF-β/Activin and Dpp/BMP subfamilies, and distant members (Newfeld et al., 1999). The TGF-β/Activin subfamily includes TGF-β1, TGF-β2, activin A, activin B, GDF8 (growth and differentiation factor 8), and so on. The Dpp/BMP family is the largest subfamily, which consists of 60A/BMP clusters, Dpp/BMP clusters, and other family members e.g. GDF3, BMP3, GDF5, etc. The *C. elegans* genome encodes four TGF-β homologs DBL-1, DAF-7, UNC-129, and TIG-2. Three of them belong to Dpp/BMP subfamily, while DAF-7 shares similarity with TGF-β/Activin family. However, the DAF-7 bootstrap value in TGF-β/Activin family is extremely low. Furthermore, phylogenetic analysis of type I and type II receptors and Smad proteins showed that all are closer to BMP family instead of Activin. Consequently, all *C. elegans* TGF-β signal pathways function similarly to BMP subfamily.

TGF-β isoforms are synthesized as TGF-β precursors containing a C-terminal TGF-β sequence and an extra N-terminal peptide (propeptide). TGF-β is cleaved from the precursor but still attached to it noncovalently before secretion (Derynck et al., 1985; Oklu and Hesketh, 2000). Mature TGF-β associates with propeptide and latent TGF-β binding protein through disulfate bond, which prevents TGF-β interaction with receptors.

In addition, inhibitory proteins exist in order to regulate the signaling pathway. For example, noggin and chordin are BMP antagonists, which bind to BMP4 and prevent it from interacting with TGF-β receptors. Inhibin α chain competes with Activin in binding to transmembrane receptors. Follistatin inhibits Activin or BMP-7 signaling

through physical association with the ligands (Massague and Chen, 2000; Balemans and Van Hul, 2002).

2. TGF-β Receptors

Three kinds of TGF-β receptors were characterized by ligand cross-linking methods (Cheifetz et al., 1987): type I receptor, type II receptor, and type III receptor. Type I and type II receptors mediate TGF-β signaling pathway; while type III receptor facilitates recruiting ligand to type II receptor in TGF-β1 and TGF-β2 signaling (Massague, 1998; Chen et al., 2003).

Both TGF-β type I and type II receptors are N-glycosylated glycoproteins containing a short extracellular region, a transmembrane helix domain, and a cytoplasmic serine/threonine kinase domain (Wells et al., 1997). Type II receptor is able to be autophosphorylated with or without ligand association, but the phosphorylation is required for TGF-β signaling activation (Luo and Lodish 1997). Type I receptor contains a GS domain, which has a SGSGSG sequence, preceding the kinase domain. Activated-type II receptor complex phosphorylates type I receptor GS domain (Wrana et al., 1994; Huse et al., 1999), which then leads to phosphorylation of type I receptor kinase domain. This activated ligand-receptor complex initiates signal transduction.

Type I and type II receptors exist as homodimers before association with ligand. The affinity of interaction of TGF-β with the two receptors is different. For example, TGF-β1, TGF-β2, and activins associate with type II receptor in absence of type I receptor; whereas BMP-2, BMP-4, and BMP-7 interact with type I receptor without type

II receptor. However some ligands, e.g. TGF-β2, has high affinity of binding to type II-type I receptor complex (Derynck and Feng, 1997).

Combinations of different receptors exist, resulting in complex cell responses. For example, TβR II binds to TβR I, which leads to activation of Smad2 and Smad3. Alternatively, TβR II can associate with ALK1 resulting in activation of Smad1 and Smad5 (Goumans et al., 2002).

Type III receptors include betaglycan and endoglin, whose major function is to recruit ligand to type I and type II receptors. However, they do not widely function in all TGF-β signaling pathways, only in TGF-β1, TGF-β2 and TGF-β3 pathway (Massague, 1998).

3. Smads

Smad proteins are critical components in TGF-β signaling pathways. The name results from a combination of Mad in *Drosophila* (Sekelsky et al., 1995; Raftery et al., 1995) and Sma proteins in *C. elegans* (Savage et al., 1996). Smad proteins are categorized into three major classes based on their structure and function: R-Smad: receptor –regulated Smads (e.g. Mad, Smad2, Smad3, SMA-2, SMA-3); Co-Smad: common mediator Smads (e.g. Medea, Smad-4, SMA-4); and I-Smad: inhibitory Smads (e.g. Dad, Smad-6, Smad-7).

Smads are transcription factors that share high homology at both N-terminal sequence (Mad Homology 1/MH1 domain) and C-terminal sequences (Mad Homology 2/MH2 domain). The MH1 domain contributes to DNA binding (Massague, 1998); while the MH2 domain regulates protein-protein interaction. For example, the MH2 domain

mediates R-Smad association with receptors and with Co-Smads (Macias-Silva et al., 1996); it also mediates formation of homomeric and heteromeric Smad complexes (Shi and Massague, 2003; Wu et al., 2002). Between MH1 and MH2 is a linker region, which in some Smads is phosphorylated by MAP kinase (Kretzschmar et al., 1997). R-Smad MH2 domain specifically contains a SXS motif in the C-terminal tail that can be phosphorylated by type I receptor upon signal activation. This motif is absent in I-Smads and Co-Smads. In contrast to R- and Co- Smad, I-Smads only have MH2 domain but do not contain MH1 domain.

Two models explain I-Smad mediated signaling antagonism: on the one hand, I-Smads compete with R-Smads for binding to receptors; on the other hand, I-Smads recruit E3-ubiquitin ligases Smurf1 (Smad-ubiquitination-regulatory factor 1) and Smurf2 that leads to receptor ubiquitination and degradation (Shi and Massague, 2003).

Mechanism of regulation occurring at the R-Smad level include auxiliary proteins for R-Smad recognition by receptor, and ubiquitin-mediated degradation via Smurf protein. The auxiliary proteins include SARA (Smad anchor for receptor activation), Axin, Dab-2, TRAP-1, Hgs/Hrs, which interact with type I or type II receptor and R-Smads. These adaptor proteins were proposed to facilitate TGF-β-induced receptor internalization, R-Smad recruitment to receptors, and vesicle trafficking (Derynck and Zhang, 2003). For example, SARA associates with Smad2/Smad3 and is localized to early endosoms during the inactive state. Upon TGF-β signaling activation, SARA bound Smad2/Smad3 are phosphorylated. However, the phosphorylation is much more efficient at SARA-rich endosomes than at the plasma membrane because receptors are internalized

into endosome through clathrin coated-pits (Di Guglielmo et al., 2003; Hayes et al., 2002).

Besides auxiliary proteins, regulation occurs through Smad degradation. The E3 ligases Smurf1, Smurf2, and SCF/Rac1 have been reported to mediate R-Smad and Co-Smad degradation with or without signaling activation (Derynck and Zhang, 2003; Izzi and Attisano, 2004). In addition, Smurf1 and Smurf2, recruited by R- Smads or I-Smads, can target other components in TGF-β signaling pathways including TGF-β receptors and the transcriptional repressor SnoN (Izzi and Attisano, 2004).

4. Gene Transcription

A Smad MH2 domain fused with GAL-4 DNA binding domain showed that a Smad complex can directly activate gene transcription (Liu et al., 1996). This was the first evidence that Smad proteins have transcription activity. Further studies showed that the Smad MH1 domain recognizes and binds to the DNA sequence CAGAC through a conserved structural element (Dennler et al., 1998; Zawel et al., 1998; Shi et al., 1998). Thus, the Smad complex has basal transcription activity. However, since the Smad – DNA binding affinity is low and the binding sequence has low complexity, Smad complexes require other transcription factors to facilitate signal specificity in different cells, tissues, and organisms (Shi et al., 1998; Massague, 1998; Derynck and Zhang, 2003). Many transcriptional activators, transcriptional cofactors, and transcription repressors were demonstrated to facilitate the Smad complex gene transcription, which is discussed below.

4.1. Transcriptional activators

FAST-1, a DNA binding protein belonging to winged-helix family, has been shown to interact with Smad2-Smad-4 and Smad3-Smad4 complexes as a transcriptional cofactor in the TGF-β/Activin but not the BMP pathway (Chen et al., 1996; Zhou et al., 1998). FAST-1 is necessary to establish the panel of homebox genes in the specification of mesoderm (Watanabe and Whitman, 1999). In the BMP pathway, OAZ, a multiple zinc finger motif protein, functions as a DNA binding cofactor of Smad1-Smad4 complex (Hata et al., 2000). Neither FAST-1 nor OAZ contains transactivation activity. Rather, they help Smads to define DNA binding specificity, and it is the Smad complex that has transcription activity.

Basic helix-loop-helix (bHLH) leucine zipper transcription factor TFE3 interacts with coactivator p300/CBP and E-box on the *PAI-1* promoter. Upon TGF-β signaling, this complex allows TFE3 to recruit Smad3-Smad4 to promote specific gene transcription (Hua et al., 1999). Transcription factor AP-1/bZIP family members, e.g. c-Jun and c-Fos, interact with Smad3 directly, to synergize in TGF-β induced gene transcription (Zhang et al., 1998).

Mammalian AMLs (acute myelogenous leukemia), also known as core binding factor (CBFS)/ polyoma enhancer binding protein (PEPB2s), have been shown to interact with R- Smads and Smad4 (Ito and Miyazono, 2003). Most importantly endogenous RNX2 interacts with endogeneous Smad1 (Zhang et al., 2000). AML is composed of an α subunit and a β subunit. The α subunits contains a Runx DNA binding domain and has three isoforms: αA, αB, and αC, which correspond to mammalian Runx1, Runx2 and Runx3. Two homologs exist in *Drosophila: runt, lozenge;* and one in *C. elegans: rnt-1.*

Paired-like homeobox genes Mixer and Milk can directly interact with Smad2, which is required for gene transcription in TGF-β/Activin pathway (Germain et al., 2000; Randall et al., 2002). Other identified Smad transcriptional cofactors include lymphoid enhancer-binding factor 1 (LEF1), ATF-2, and so on, which enhance Smad DNA binding specificity (Derynck and Zhang, 2003; Itoh et al., 2000; Massague, 2000).

In addition, the Smad complex is able to enhance or reduce gene transcription by recruiting transcriptional coactivators or corepressors although they do not directly change DNA binding specificity or affinity of the Smad complex.

4.2. Transcriptional coactivator

p300 and CBP are two well-known transcriptional coactivators recruited by Smads (Derynck et al., 1998). Both contain multiple regions for interaction with different transcription factors and both also have histone acetyl transferase (HAT) activity that can remodel nucleosomal and chromatin structure and subsequently increase accessibility of transcriptional machinery. Other Smad transcriptional coactivators include ARC105, MSG1, SMIF (Derynck and Zhang, 2003). These transcriptional activators interact with R-Smads and Smad4 through the Smad MH2 domain.

4.3. Transcriptional Corepressors

Ski and SnoN are two well documented Smad transcriptional corepressors. Ski was first identified in transforming retroviruses isolated from chicken embryo cells (Li et al., 1986). Both are nuclear proteins and have orthologs in other organisms e.g. human, mouse, frog, zebrafish, and *C. elegans* (Liu et al., 2001; da Graca et al., 2004). Ski and

Sno have been shown to interact with R-Smads and Smad4 to inhibit TGF-β responsive gene transcription (Stroschein et al., 1999; Xu et al., 2000). Studies of Ski structure showed that Ski and phosphorylated R-Smads share a similar binding surface on Smad4 (Wu et al., 2002). Thus, it suggests that negative regulation by Ski on the TGF-β signaling pathway results from Ski interfering with the formation of Smads complex. Meanwhile, Ski can directly interact with other transcriptional corepressors such as mSin3A, HIPK2, MeCP2, N-CoR etc, to further repress gene transcription. In addition, Ski and SnoN have been shown to prevent Smads from binding to coactivator p300/CBP.

Regulation of SnoN by TGF-β occurs at both translational and transcriptional levels, but with opposite effects. Upon TGF-β signal activation, Smads recruit Smurf2 or E3 ubiquitin ligase anaphase promoting complex (APC) and target degradation of Smad bound SnoN (Bonni et al., 2001; Wan et al., 2001). Then the Smads are released in order to stimulate gene transcription. Secondly, SnoN transcription was significantly increased upon TGF-β stimulation (Wotton and Massague, 2001; Luo, 2004). However, this effect occurs two hours later, and may serve as a negative feedback effect on the TGF-β signal pathway.

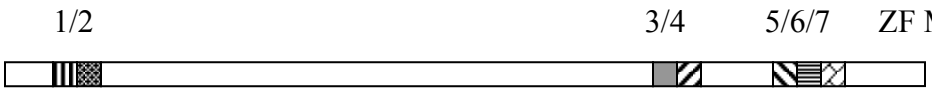
TGIF is a homeodomain containing protein and acts as a negative regulator of theTGF-β signal pathway. TGIF competes with coactivator p300/CBP for binding to Smad2 in a TGF-β-dependent manner. Meanwhile, TGIF recruits histone deacetylases (HDACs) to the Smad complex. Since HDAC leads to nucleosomal packing that would decrease transcription machinery accessibility, the HDAC-TGIF-Smad complex represses gene transcription. It has been reported that the EGF-Ras-Mek pathway can

phosphorylate TGIF, there by stabilizing this short-lived protein and repressing TGF-β responsive target gene transcription (Lo et al., 2001).

SIP1 (Smad-interacting protein 1) contains a homeodomain, a Smad-interacting region, and two zinc finger regions for binding to DNA (Remacle et al., 1999). Through association with Smad1 MH2 domain, SIP1 was removed from DNA binding so that repression of gene transcription was relieved. Other transcriptional corepressors include c-myc, Evi1, ATF3, SNIP1, Tob, Hoxc-8 (Derynck and Zhang, 2003; Shi et al., 1999; Kurokawa et al., 1998).

4.4. *schnurri* and *brinker* in the *Drosophila* Dpp pathway

In *Drosophila,* one of the most extensively studied transcription factors is Schnurri (Shn) in the Dpp pathway. *Shn* encodes a zinc finger transcription factor homologous to human MBP1/PRDII-BF1/HIV-EP1 (Arora et al., 1995; Grieder et al., 1995; Staehling-Hampton et al., 1995). Schnurri has been shown to associate with Mad and Medea as a transcriptional cofactor, which is required for Dpp signaling (Dai et al., 2000). *Shn* loss of function (lf) mutants are embryonic lethal as are other Dpp pathway component mutants. Shn contains seven C_2H_2 zinc finger (ZF) motifs, which are grouped into three clusters. Dai et al (2000) showed that the first two clusters have the ability to bind DNA elements with different sequence selectivity. The third one – a triplet zinc



Shn schematic structure

finger motif - contributes to binding the Silence Element (SE) of *brinker (brk)*.

Brinker acts as a transcriptional repressor in the Dpp pathway (Jazwinska et al., 1999). It has been demonstrated that Brinker interacts directly with Dpp-responsive gene *cis*-regulatory elements (Rushlow et al., 2001; Sivasankaran et al., 2000). In *brk* mutants many Dpp target genes were upregulated. *brk* (lf) blocks gene downregulation caused by *shn* (lf) (Marty et al., 2000). Several lines of evidence showed that Shn mediates Dpp-dependent repression of *brk* gene transcription (Affolter et al., 2001). It has been shown that C-terminal 670 aa sequence of Shn, including the triplet, are necessary and sufficient for Dpp-dependent repression of *brk.* This region can form a stable complex with Mad and Medea and SE (Muller et al., 2003).

Consequently, the *brk* expression level is opposite to the Dpp morphogen gradient. The further away from the Dpp source, the higher the expression of Brinker. Until recently, Brk was considered to play a central role in the Dpp pathway as a transcriptional repressor and the major functions of Shn had been thought to mediate repression of *brk*. Pyrowolakis et al (2004) showed that Shn mediates repression of other target genes through sites similar to the SE isolated from *brk.* Thus, it suggests that Shn mediated gene repression is solely dependent on the SE but not *brk.*

5. Smad Ubiquitination Regulatory Factor - Smurf

Smurfs are E3 ubiquitin ligases and have a HECT (homologous to E6-associated protein C terminus) domain, which include two family members Smurf1 and Smurf2. They also contain a WW domain that interacts with Smad PPXY motif in the linker region. Smurf1 was cloned in *Xenopus* (Zhu et al., 1999). This was the first

demonstration that TGF-β signaling pathway was modulated by ubiquitination. Smurf1 only interferes with the BMP pathway by targeting Smad1 and Smad5 specifically but not Smad2/3 etc (Zhu et al., 1999). Smurf-2 (Kavsak et al., 2000; Lin et al., 2000; Zhang et al., 2001) has a broader specificity that affects both BMP and activin pathways. SCF/Roc1 complex, another E3, mediates activated Smad3 ubiquitination (Fukuchi et al., 2001).

At the basal level, Smurf1 targets and degrades R-Smads (e.g. Smad1/5) in a ligand independent manner (Zhu et al., 1999). Upon TGF-β signaling activation, ubiquitination occurs at different levels. Smurf1 and Smurf2 are recruited by I-Smad (e.g. Smad6/7) and target activated type I receptor to prevent R-Smad phosphorylation (Kavsak et al., 2000; Ebisawa et al., 2001). Smurf2 targets Smad2 bound SnoN for ubiquitination through the HECT domain (Bonni et al., 2001), so that Smad2 can respond to TGF-β signaling activation.

6. Crosstalk

Crosstalk has been investigated between TGF-β and other signaling pathways. There is strong evidence for crosstalk with Erk-MAPK pathway (Kretzschmar et al., 1997), Wnt/β-Catenin pathway (Baker et al., 1999; Nishita et al., 2000), Ca^{2+}/Calmodulin signaling (Zimmerman et al., 1998; Scherer and Graff, 2000), JAK-STAT pathway (Nakashima et al., 1999), p53 network (Cordenonsi et al., 2003), other TGF-β subfamily pathways (Nishita et al., 1999), and several other pathways.

III. TGF-β Signaling pathway in *C. elegans*

The *C. elegans* genome encodes four TGF-β homologs: DBL-1, DAF-7, UNC-129, and TIG-2 (Savage-Dunn, 2001). Two major TGF-β pathways have been characterized in *C. elegans*: Dauer pathway and DBL-1 pathway. Both are more closely related to BMP/Dpp pathways than to those of other TGF-β superfamily members (Newfeld et al., 1999; Suzuki et al., 1999; Colavita et al., 1998). The function of TIG-2 and UNC-129 are still poorly understood.

1. Dauer Pathway

The TGF-β Dauer pathway is one of the pathways that regulates dauer formation (Patterson and Padgett, 2000; Shostak et al., 2004). Many signaling components of this pathway have been characterized by genetic screening and gene cloning. Under tough living conditions such as a shortage of food or hot temperature, L2 worms develop into dauers for survival. Normally, these dauers can be recovered when living conditions are improved. Dauer pathway (lf) mutants have two basic phenotypes: constitutive dauer – L2 forms dauer with sufficient food source at 25°C and the dauer is difficult to recover; or dauer defective - animal is not able to develop dauer under dauer formation condition such as 25°C or insufficient supply of food. Based on gene sequence similarity and genetic studies, a model of dauer pathway signaling was established based on other TGF-β pathways.

DAF-7 shares sequence homology to TGF-β superfamily and acts as the ligand (Ren et al., 1996). DAF-4 belongs to type II receptor family and is shared with the DBL-1 pathway (Estevez et al., 1993). DAF-1 is the type I receptor that phosphorylates two

Smad proteins: DAF-8 and DAF-14 (Georgi et al., 1990). Unlike other Smads, DAF-8 MH1 domain is not highly conserved; DAF-14 even excludes the MH1 domain. One interesting phenomenon is that the type I receptor DAF-1 functions in the absence of type II receptor DAF-4 in the Dauer pathway (Gunther et al., 2000). However, like other TGF-β pathways the type II receptor is unable to function independently of ligand activation, although it does contain a self activated kinase activity. Loss of function of all these genes - *daf-7, daf-4, daf-1, daf-8, daf-14* - gives rise to dauer constitutive phenotype.

The gene *daf-3, daf-5,* and *daf-12* function downstream of the TGF-β dauer pathway. *daf-*3 encodes a Smad protein with MH1 and MH2 domains, but lacks a C-terminal SSXS phosphorylation site (Patterson et al., 1997). *daf-5* encodes *C. elegans* Ski (da Graca et al., 2004). DAF-3 and DAF-5 directly interact with each other. Mutants of *daf-3*, *daf-5*, and *daf-12* are dauer defective. Genetic studies showed that either *daf-3* or *daf-5* mutations suppress the *daf7/4/1/8/14* dauer constitutive phenotype (Patterson and Padgett, 2000), suggesting that *daf-3* and *daf-5* are downstream components. It also suggests that unlike Ski/Sno protein, DAF-5 acts as a cofactor to Smad protein DAF-3 instead of as an antagonist (da Graca et al., 2004).

The downstream component *daf-12* encodes a nuclear hormone receptor (Antebi et al., 2000) that has a high homology to vertebrate vitamin D receptor. It has been proposed that the TGF-β and insulin pathways converge at *daf-12* in regulating dauer formation (Shostak et al., 2004; Gerisch et al., 2001).

Other proteins were demonstrated to function in this pathway. For example, BRA-1 has high homology to BRAM1 (BMP receptor associated molecule 1) protein and was shown to interact with DAF-1 *in vitro*. Double mutants of *bra-1* and *daf-1/7* suppress

dauer formation significantly, which suggests that BRA-1 negatively regulates the Dauer pathway through association with the receptor (Morita et al., 2001).

2. DBL-1 Pathway

The DBL-1 pathway was first characterized by Savage C et al (1996). The group cloned three Smads: R-Smads *sma-2*, *sma-3* and Co-Smad *sma-4*, and characterized the pathway phenotype: small body size (Sma) and abnormal male tail fusion (Mab). Later on, other related genes were cloned, e.g. *dbl-1* as ligand (Suzuki et al., 1999), *sma-6* as type I receptor (Krishna et al., 1999) etc, *lon-1* (Maduzia et al., 2002; Morita et al., 2002), *lon-3* (Nystrom et al., 2002; Suzuki et al., 2002), *rnt-1*(Nam et al., 2002), *kin-29* (Maduzia et al., 2005) and *sma-9* (Liang et al., 2003). This section focuses on functional regulation of the DBL-1 pathway instead of signaling components.

2.1. Body Size

All DBL-1 pathway mutants (both hermaphrodite and male) have small body size compared with wild type, for example, *daf-4, dbl-1, sma-3,* and so on (Savage-Dunn et al., 2000). In principle, body size may be determined by either cell number or cell size. Nuclear counts of Sma mutants demonstrated that they have the same cell number as wild type animals (Suzuki et al., 1999). Furthermore, Rafal Tokarz (our former lab member) measured animal seam cell size and L3 worm length and showed that the Sma mutant seam cells are shorter than the wild type ones (Fig 5B). This suggests that it is cell size not cell number that causes small body size of DBL-1 pathway (lf) mutants.

The DBL-1 pathway regulates body size development post-embryonically. Worm length measurements showed that the mutant animal shares the same body length at the L1 stage as wild type. However they grow slowly so that the adult is only 50% the length of wild type (Fig 5A). Expression patterns of *dbl-1* and *daf-4* further support this idea: *dbl-1* transcription was absent in early embryos and displayed in late embryo, larval stages and adults (Suzuki et al., 1999). *daf-4* expressed from L1 to adult (Gunther et al., 2000). Over-expression of *dbl-1* in wild type background gives a long (Lon) phenotype. The animal increases its body length and gives rise to a specific male ray fusion defect (Suzuki et al., 1999). This suggests that ligand DBL-1 functions in a dose-dependent manner in body size development.

dbl-1 is primarily expressed in neurons - ventral nerve cord, pharyngeal neurons, and male glial cells (Suzuki et al., 1999). The type II receptor gene *Daf-4* is expressed in the pharynx, intestine, hypodermis, body wall muscles, and head neurons of the lateral, vesicular and retrovesicular ganglia (Gunther et al., 2000). Both *sma-6* and *sma-3* are expressed in the pharynx, intestine, and hypodermis (Wang et al., 2002; Krishna et al., 1999), but not in neurons. All of the components are expressed post-embryonically, except *sma-3* which is transcribed in embryonic and larval stages and in adults. Since the pathway needs other components for activation, these embryonic *sma-3* RNAs may not lead to active SMA-3 until after L1.

The downstream component *lon-1* is expressed in hypodermal and intestine cells post embryonically. *lon-3* is expressed in most hypodermal cells on the surface of larval and adult animals. Expression of *daf-4*, *sma-6, sma-3,* and *lon-1* in hyp7 cell alone is

sufficient to rescue comparative mutant body length (Gumienny and Padgett, 2003). Thus, hyp7 is a critical tissue for DBL-1 pathway regulated body size development.

Unlike other pathway components, *dbl-1* was mostly expressed in neuronal cells. The question arises as to how does the signal of *dbl-1* is transduced by the downstream components since they do not locate to the same tissue. The current model supports the idea that DBL-1 might be synthesized and secreted from these neurons and then diffuse into the target tissues e.g. hyp7 cell for regulating body size.

Further investigation of DBL-1 pathway (lf) related Lon and Sma phenotypes showed a direct correlation between the DNA content in hyperploid hypodermal nuclei and body size (Flemming et al., 2000; Morita et al., 2002). As hyp7 nuclei undergo endoreplication during larval development, they become polyploid, and the ploidy in these nuclei is about 12 in wild type animals. The authors showed that the Sma mutant has a decreased ploidy, while the Lon mutant has an increased ploidy. These studies indicate that the DBL-1 pathway controls endoreplication in hyp7 in order to regulate body size.

Downstream regulators

The negative regulators *lon-1* and *lon-3* may function as downstream components in the DBL-1 pathway. Loss of either *lon-1* or *lon-3* function gives rise to a Lon phenotype. Both genes affect body size development in late larval stages but do not regulate male tail morphogenesis. *lon-1* encodes a cysteine-rich secretory protein (CRISP), whose transcription is negatively regulated by the DBL-1 pathway (Maduzia et al., 2002).

It has been reported that germ-line signals regulate body size (Patel et al., 2002). Laser ablation of germ-line precursor cells (Z2 and Z3) gives rise to a giant adult without a germline. These animals are larger in volume and longer in length than wild type animals. However, the phenotype is different from DBL-1 pathway caused Lon phenotype, since these mutants do not increase animal volume but only the length (Hirose et al., 2003).

lon-3 encodes a cuticle collagen and is regulated post-transcriptionally by the DBL-1 pathway. Based on the position of cysteines, LON-3 belongs to the SQT-1 subgroup of collagens (Suzuki et al., 2002). Collagen genes *rol-6* and *sqt-1* are required for *dbl-1* regulated body length and genetically interact with each other for animal morphology (Kramer et al., 1990). It has been reported that *rol-6* and *sqt-1* genetically interact with *lon-3,* and these interactions are reciprocal (Nystrom et al., 2002). Loss of *rol-6* or *sqt-1* function hardly changes the animal body size. However, either *rol-6 (lf)* or *sqt-1 (lf)* can suppress *lon-3(lf)* and *dbl-1*(overexpression) phenotype. These results support that collagen genes play an important role in DBL-1 pathway regulated body size development.

kin-29 encodes a serine-threonine kinase related the EMK (ELKL motif kinase) family, which interacts with microtubules. Loss of *kin-29* function results in small body size, but does not cause a male tail defect. *kin-29* expression in neurons or in the hypodermis rescued mutant's small body size. Genetic studies placed *kin-29* downstream of *dbl-1* and upstream of *lon-1* (Maduzia et al., 2005).

BRA-2 shares high homology to BRAM1 and BRA-1 (Morita et al., 2001). It was expressed in the pharyngeal muscle and intestine, which overlaps some of the expression

pattern of SMA-6, indicating that BRA-2 might act as a receptor-associated factor in the DBL-1 pathway. BIP (BRAM interacting protein) interacts with BRA-1 and BRA-2 *in vitro* (Sugawara et al., 2001). It is expressed in pharyngeal muscle, hypodermis and several neurons. BIP dsRNAi caused a Sma phenotype. Also BIP dsRNAi suppressed the Lon phenotype caused by dbl-1 over-expression, but not by *lon-1(lf)*. Thus, BIP might be a regulator of body size development via interacting with BRA-2.

rnt-1 encodes the only homolog of mammalian *AML/RUNX* genes in *C. elegans* (Nam et al., 2002). *rnt-1* translational expression pattern showed that it is expressed in seam cells from embryonic bean stage to L3 stage, in intestine cells from late embryonic stage to L3 stage, and the V5, V6, and T descendant cells in the male tails. Loss of function mutants are shorter than wild type but slightly longer than Sma mutants, and the male tails have a broad range of defects in ray1 to ray9 but do not have a defect in the spicules. Genetic studies put *rnt-1* upstream of *lon-1*. *In vitro* GST pulldown assay showed that RNT-1 interacts with SMA-4. However, double mutants of *rnt-1* and *sma-2, sma-3, sma-4,* and *sma-6* showed a much shorter body size and more severe male ray fusion than single mutants, suggesting that *rnt-1* and DBL-1 pathway components function in a synergistic manner for regulating body size and male tail morphogenesis (Ji et al., 2004). It is still unclear whether or not *rnt-1* functions in the DBL-1 pathway.

2.2 Male Tail Morphogenesis

The male sensory rays and spicules are critical for mating as described above. Examination of *daf-4, sma-2, sma-3, sma-4, and sma-6* mutant males demonstrated abnormal ray fusions and crumpled spicules. The fusions frequently occurred at rays 4-5,

6-7, and 8-9 (Savage et al., 1996). The mutant male spicules are twisted and short with high penetrance instead of long and sharp as in wild type. As a result, these animals fail to mate sucessfully.

The sensory ray defects of DBL-1 pathway (lf) mutants are due to defects in sensory ray identity (Lints and Emmons, 1999). Two neurons - RNA and RNB – are present in each male ray. Dopamine (DA) is expressed in the R5A, R7A, and R9A neurons in wild type animals. Adoption of the dopaminergic fate is accompanied by tyrosine hydroxylase activity. CAT-2 encodes tyrosine hydroxylase (TH) that catalyzes the first rate-limiting step of DA biosynthesis. In wild type cells, the frequency of CAT-2::GFP expression in R5A, R7A, and R9A is over 95%. However, in *dbl-1, sma-6, sma-4,* and *daf-4* mutants, it drops to 10-20% for R5A, 20-40% for R7A, and 0-2% for R9A cells. Non-sex-specific dopaminergic neurons were not influenced.

The Hox gene *egl-5* regulates male ray development and is necessary for expression of DA in R5A neuron. A strong genetic interaction between *egl-5* and DBL-1 pathway has been reported: *egl-5* is necessary for the DBL-1 pathway-dependent expression of DA by rays descended from the blast cell V6 (Lints and Emmons, 1999). *mab-23* encodes a DM (Doublesex/MAB-3) domain transcription factor and is required for male DA fate identity. This suggests that *mab-23* interacts with *egl-5* and the DBL-1 pathway in stimulation of DA fate specifically in ray5A.

lin-31 encodes a forkhead/winged helix protein. The mutant has crumpled spicules and defects in migration of the tail cell that leads to an abnormal cellular mold for the spicules (Baird and Ellazar, 1999). Since the forkhead/winged helix protein FAST-1 acts as a transcriptional cofactor in the TGF-β/activin pathway, this suggests that

LIN-31 may be a cofactor in the DBL-1 pathway regulating male tail morphogenesis, in particular the spicules.

2.3 Innate Immunity

DBL-1 pathway is one of the signaling pathways that regulates *C.elegans* innate immunity (Nicholas and Hodgkin, 2004). Studies of bacterial infection on *dbl-1(lf)* mutants showed that the animals were hypersensitive to the pathogen (Millet and Ewbank, 2004; Kurz and Tan, 2004). Meanwhile, *dbl-1* mutants lived shorter than wild type when using the bacterial strain OP50 as a food source (Mallo et al., 2002). All these phenomena support that the DBL-1 pathway regulates innate immunity in *C. elegans.* The intestine is the organ for digestion and direct contact of bacterial pathogens. As mentioned above, DBL-1 pathway components have strong expression in the intestine where they might regulate target genes in response to pathogens. The mechanism of how the DBL-1 pathway controls *C. elegans* innate immunity still remains unclear.

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

Our lab uses *C. elegans* as a model organism to study DBL-1 pathway regulated body size development and male tail morphogenesis. In order to isolate additional DBL-1 pathway components, genetic screens were performed by Dr. Cathy Savage-Dunn (Savage-Dunn et al., 2003). One of the mutants isolated is *sma-9,* which shows the similar body size and male tail phenotype as other DBL-1 pathway components. Thus, it suggests that *sma-9* may be function in the DBL-1 pathway. To address this issue, I started to clone this gene and study its physiological function in animal development.

Chapter 1. The *Caenorhabditis elegans schnurri* Homolog, *sma-9,* Mediates Stage- and Cell Type- Specific Responses to DBL-1 BMP-related Signaling

ABSTRACT

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, *sma-9* loss of function phenotype, and the expression pattern support that *sma-9* acts as a downstream component and is required in the DBL-1 signaling pathway, providing the first evidence for 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.

INTRODUCTION

The bone morphogenetic proteins (BMPs) are members of the transforming growth factor β (TGFβ) superfamily of secreted peptide growth factors that play critical 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 suggesting that, in vivo, efficient promoter binding and target gene regulation may require interaction with transcriptional cofactors (Shi et al., 1998; Massague 1998). To date, however, few Smad-interacting transcriptional cofactors have been identified: *Drosophila* Schnurri (Shn), and vertebrate FAST-1, FAST-2, OAZ, and Mix (Arora et al., 1995; Grieder et al., 1995; Staehling-Hampton et al., 1995; Chen 1996; Hata 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 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 possibility 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 Shn-1 (MBP-1/PRDII-BF1/αA-CRYBP1/GAAP-1/HIV-EP1), Shn-2 (MBP-2/HIV-EP2), and Shn-3 (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, 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 that 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. Emmons, 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); *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 noncomplementation 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 nonUnc animals. From a screen of approximately 4000 F1 cross progeny, 9 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). 20ng/μl cosmid DNA was injected into *sma-9(wk55)* for rescue. 10ng/μl plasmid DNA of GFP constructs was injected into N2.

Sequencing *sma-9* mutants

15 mutant animals were picked into 10μl lysis buffer (10mM pH8.0 Tris; 50mM KCl; 2.5mM $MgCl_2$; 0.45% Tween 20; 0.01% gelatin; 60μg/ml proteinase K) and placed at -80°C for one hour. The frozen solution was heated to 60°C for 1 hour and then 20 min at 95°C to generate crude lysate. PCR was carried out on mutant and wild type genomic DNA templates, using platinum *Taq* and platinum *Pfx* mixture as DNA polymerase, and primers within genomic sequence. For *wk55,* the region from 5532 to 14827 in T05A10 covering the whole open reading frame (ORF) have been sequenced. For *qc3*, the regions from 5532 to 7558 and 11151 to 13081 in T05A10 have been 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 are: *yk1285a11, yk128a8, yk1136g02, yk1109f01, yk43h3, yk6d10, yk864c1, yk1134e06, yk856b10, yk1057a6, yk1216e10, yk1237d01, yk1103h10***,**

yk127d10, yk1264e07, yk328c9, yk228h6. All yk cDNAs are gifts from Y Kohara. Details for all the primers used for sequencing all the yk clones are available upon request.

LiCl RNA Preparation from wild type animals was performed as described at http://www.dartmouth.edu/artsci/bio/ambros/protocols.html (method from Ambros Lab). RT-PCR was performed using SUPERSCRIPT^{TM} One-Step RT-PCR (Invitrogen). pCS234 (cloned into pBluescript SK+ at *Not*I and *Bam*HI) starts from the predicted exon 1 (5' variant A in Fig. 3), splices out exon 4, exon 5 and ends at exon 9 overlapping with *yk328c9*. Primers used for cloning pCS234 are: AAGCGGCCGCATGAGCCATCAGGCAATTGG and AAGGATCCGGTTCAAGGTTTTGTGTCAC. pCS272 (cloned into pBluescript SK+ at *Sma*I and *Bam*HI) starts from exon 4 and ends at exon 9 overlapping both *yk1285a11* and *yk328c9*. Primers used for cloning pCS272 are: AACCCGGGCCCCTCGCTCTCCAAA and AAGGATCCTGATGGTCCTTG.

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

2.8kb: GCGCGGCCGCAAACATTTGTGAAGTTG; AAGGTACCTTCGCCAATTCTAAAACCACT.

5.5kb: AAGCGGCCGCGAGTTCACACAGTTTATGAT;

AAGGTACCTTCGCCAATTCTAAAACCACT.

8.0kb: AAGCGGCCGCCCATCCAATATTCAATTCTT;

AAGGTACCTTCGCCAATTCTAAAACCACT.

1.5kb: GGCCGCGGAGTTCACACAGTTTATGAT; GGGCGGCCGCCGAAAATTGCAGGTCTG.

4.0kb: GGCCGCGGCCCATCCAATATTCAATTCTTTA; GGGCGGCCGCCGAAAATTGCAGGTCTG.

The 2.8kb fragment was cloned into pPD117.01 at *Not*I and *Kpn*I (pCS231), then into pBluescript SK+ (*Bam*HI and *Kpn*I), finally into pPD117.01 at *Xba*I and *Kpn*I (pCS251).

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

dsRNAi

The templates used were *yk1285a11* cDNA clone (containing sequences from predicted exons 1-7 and alternative exon 1 (Figure 3)), and *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 EcoRI and KpnI; *yk228h6* 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. 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 min. 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 10min and then at 37°C for 30min. The dsRNA was microinjected directly into N2 animals without further treatment. Males with small body size were picked to score the Mab phenotype.

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 critical 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). Only viable mutants were selected in these screens, so we wished to determine whether null alleles of *sma-9* cause more severe defects or lethality. Since *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 (Materials and Methods), but none of them showed lethality or more severe defects than previously existing alleles. Thus, the existing 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, 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 analysis has been done 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) (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; *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 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 (Materials and Methods) phenocopies *sma-9* mutants, causing small body size

Fig. 1. *sma-9* mutants display Sma and Mab phenotypes similar to those of DBL-1 pathway mutants (from Liang et al., 2003)

(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). All animals in C-J are young adults photographed at the same magnification, with anterior oriented to the left.

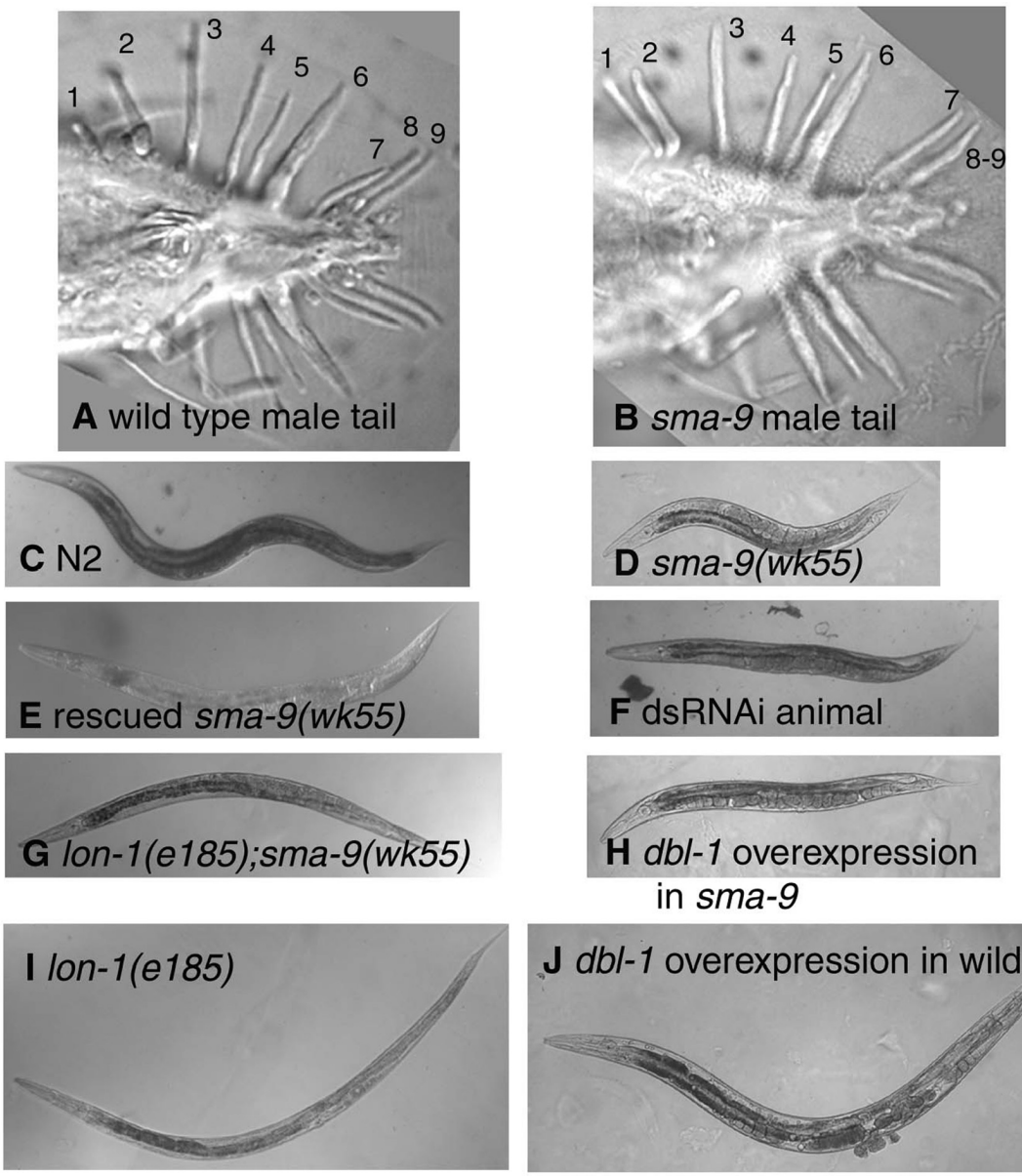

1
2
3
4
5
6
7
8 9
A wild type male tail
1
2
3
4
5
6
7
8-9
B sma-9 male tail
C N2
D sma-9(wk55)
E rescued sma-9(wk55)
F dsRNAi animal
G lon-1(e185);sma-9(wk55)
H dbl-1 overexpression in sma-9
I lon-1(e185)
J dbl-1 overexpression in wild type

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

Mutant	Fusion of Rays 4-5	Fusion of Rays 6-7	Fusion of Rays 8-9	N (sides)	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	
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	

All strains contain *him-5(e1490)*. *qcEx49* contains cosmid T05A10 and plasmid pRF4.

in both sexes and fusion between male rays 8 and 9 but no additional phenotypes (Fig. 1; Table 1). Therefore *sma-9* corresponds to the predicted gene T05A10.1.

The longest conceptual SMA-9 sequence, based on the isolated cDNA clones discussed below, is 2170 amino acids (aa) in length containing 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 to a zinc finger transcription factor family that includes *Drosophila* Shn (BLAST E value = 4e-24; Arora et al., 1995), and vertebrate Shn-1 family members human major histocompatibility complex-binding protein 1 (MBP-1)/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, S/TPKK motifs as well as stretches of sequence 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 MBP-1/

Fig. 2. *sma-9* encodes a zinc finger transcription factor (from Liang et al., 2003)

(A) The longest conceptual SMA-9 sequence. Mutation sites in *wk55* and *qc3* alleles are boxed. Seven C_2H_2 Zinc finger motifs are shaded. NLSs are underlined. S/TPXK/R or SPKK motifs are double underlined. ARDs are bold and underlined. The star 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 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 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 be 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*.

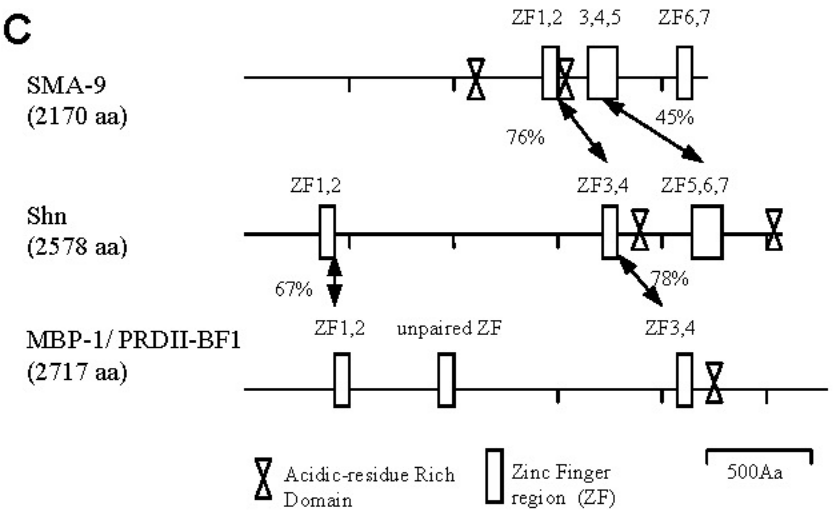
A

```
MAFAPGSGGGGAPPGQQPAQMMMPQNISAETYHRLN***MSHQAIGISNNFQ***QVQREQLNHQR  60
LLQAQLQTNGPGSVSQQQQASQQQQQVQHVQQQQQSQQQQQQQVASQQNQPQLQMNAQILQ  120
ALSTPQGQNLVNMLLMQQALNAQQTDTQNPQQIMQHQLAQQQAQQAQQAQQAQQARQQAE  180
QQAQAQARHQAEQQAQQQAQQQAQQARQQEQQAQLAAIQQQVTPQQFAQILHMQQQLQQQQ  240
FQQQQLQQQQLQQQQLQQQQLQQQQLQQQQLQQQVLQISQAQQQAQQAQHVQSRQMQPSQQ  300
SQVQAQLQQQQQLQQQQAQQLSQQQAQQQQQLQQLQLQQFLQQQQQLHQQRAAAQQAQAQ  360
NNASQQRPSVASTPALSSTPQLNDLTQTMQAQLQQQLLLQQQQAQAQQAQQAQQAQLAQQ  420
AQQQQQGQSQNRTVSQALQYIQSMQLQQRADGTPNAPSLSKPLEQPSSSKAASSGNESMS  480
DHISRIISENEVILQGDPVIRKKRPYHRQIGAQSSVDHDSNSGGSTRTSPGPKDSRMLQA  540
ASRSQSLFELSGSKHFMGSLTSGQPLLRPIQAHNDPNYTPECIYCKLTFPNEAGLQAHEV  600
VCGKKKELEKAQIAQEGNPHSALKRRHTHQDATLAMHSPLAAHTPSNMPGPSEPAIKLKK  660
DDSTELDGTSKPDALQSSSSFPRSLPKEWEQHMLTLQNLAAIIPPFVQAFLAKVLQTKLT  720
MVSSGTLTHIYENYNISPICVKQFLHFASQLTDQQLEEMTVESEKQYLENAEEYQAKGII  780
QALPCDNLEVTEMLKQQETIMKGQVPDLLVSTQAVLHHCSGKRDNKDPSKNFVMIKLLHA  840
NGKDITEIPMKAETDLDECNARFVYQMREFQKINNMNDRLIEVLKTDPAAAATPLFQVAR  900
EQSASVKFLIRQTGMVHLTIVASVAKRLGMSSMQQDVPESNQIPNGIIAFDGVQIKQ***EPE***  960
***DPPGDDDDDDDDCIIE***VVDEDQKQHIAMLTAAASGQDIGIGGQFQKVTQNLEPAIVNQQS 1020
IVNHGDSIVNNASIPIANNAPVQIVQNGGPMQHEVVAMPVQLKLEDLRLEPQTSGDTDKP 1080
YWLVINGDIGGRPSFMTTAGMTSRTHRTRNITSETYVTVPRQQPMFAVQDGTLSMYAKWN 1140
VPVHNDAETKMNLSFMGMVSLRFRTGQQKFFKYTTANKDQGHYRMTHSSFWDISTKIRDR 1200
QASMSEEKTPEESVDYDAQFIERLVGGTYTNTDLQGPSNAPVTIPVLVAPEDSTSAEPST 1260
SGQSLLMRSPRPQSPPLRDIKMDLSDDDYSATDLATTCKLEPKQESFEDVKDVKRENSPD 1320
ARPTVIISDDAGRIRRERFANKYISRIRPKHDQIIGGHRTDEVYVYVRGRGRGRYICDRC 1380
GIRCKKPSMLKKHIKSHTDVRAFNCTACNFSFKTKGNLTKHLSSKTHQRRISNIQAGNDS 1440
DGTTPSTSSMMNMDDGYHRNQPLF***DDYDNNSSDEEDYD***HLNRMQAEHKFKFGQEHILFER 1500
TAHTPPTRWCLVEAQNDHYWPSPDRRSCMSAPPVAMQRDFDDRAMTPVSGANSPYLSQQV 1560
HSPMSTSSQSNIILDIPNNQKSNCSSVSNVSPSNSQNFQSLSTVPTCASSSSNVLVPNVN 1620
FLQKDETLKCDQCDRTFRKISDLTLHQHTHNIEMQQSKNRMYQCSECKIPIRTKAQLQKH 1680
LERNHGVHMDESVTACIDPLASTQSVLGGPSTSNPRSFMCVDCDIGFRKHGILAKHLRSK 1740
THVMKLESLQRLPVDTLSLITKKDNGACLNDIDTTDCEKARISLLAIVEKLRNEADKDEQ 1800
GSVVPTTSIPAPQPVALTPEMIRALANAQTPVTASMTNTPSTAQFPVGVVSTPS*VSAVSA 1860
SGSQSNVSCVSSFNNSTMSPNPTVVPQVFPTPNPSSPLESSSMQFRKKAVLDSATHANDM 1920
PRTILRISEIPSSLPVNHQLHRDLSFLAHTTSRSESSITSPIVSSSTNFSYRKRSESSLS 1980
GSSPTHTKKLMVWNPPLAEPSFYSPKAALHPLSTDKAHASESLSDRLHNKRPRPIPDNTK 2040
CQICADEFSTPIELQVHLHVDHVRMMDGAEYKCPRKFCGLNYESLDSLRAHVTAHYETDR 2100
QKLLEEKVLLAEADFPIDNSKIEKLNSPKKESMNKFTTPFKAISDHHELYAQTQQGAGSS 2160
TSNQSPKAAN                                                     2170

LQQLHYVAQSNCGSTGVSNTKSVIPIGIVLNAVQKESSFGLGNTRERHAQNNSTNLGDSV
QSTSQSSTSS
```

B

```
                                First Zinc Finger motif
SMA-9    IIGGHRTDEVYVYVRGRGRGRYICDRCGIRCKKPSMLKKHIKSHTDVRAF
Shn      LVGGYESHEDYTYIRGRGRGRYVCSECGIRCKKPSMLKKHIRTHTDVRPF
MBP-1    FDGGYKSNEEYVYIRGRGRGKYICEECGIRCKKPSMLKKHIRTHTDVRPY
αA-CBP1  FDGGYKSNEDYVYVRGRGRGKYICEECGIRCKKPSMLKKHIRTHTDVRPY

                          Second Zinc Finger motif
SMA-9            NCTACNFSFKTKGNLTKHLSSKTHQRRISNIQ
Shn              TCSHCNFSFKTKGNLTKHMQSKTHFKKCIELG
MBP-1            HCTYCNFSFKTKGNLTKHMKSKAHSKKCVDLG
αA-CBP1          HCSYCNFSFKTKGNLTKHMKSKAHSKKCVDLG

                 Third Zinc Finger motif                    Fourth Zinc Finger
SMA-9    QKDETLKCDQCDRTFRKISDLTLHQHTHNIEMQQSKNRMYQCSECKIPIRTKAQLQ
Shn      AGDARPTCTMCSKTFQRQHQLTLHMNIHYME------RKFKCEPCSISFRTQGHLQ

                 motif                                      Fifth Zinc Finger
SMA-9    KHLERNHGVHMDESVTACIDPLASTQSVLGGPSTSNPRSFMCVDCDIGFRKHGILA
Shn      KH-ERSEA-HKNKVMMT---------STFGVPTTSNPRPFECTDCKIAFRIHGHLA

                 motif
SMA-9    KHLRSKTHVMKLESLQRLPVDT
Shn      KHLRSKTHVQKLECLQKLPFGT
```

C



PRDII-BF1, and 74% to the second pair in αA-CRYBP1 (Fig. 2B). Therefore, SMA-9, MBP-1, αA-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 MBP-1 and αA-CRYBP1 (Fig. 2B), suggesting its elimination during vertebrate evolution or 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 its unique function in *C. elegans*. An alternative 70 aa C-terminus is also unique to *C. elegans*.

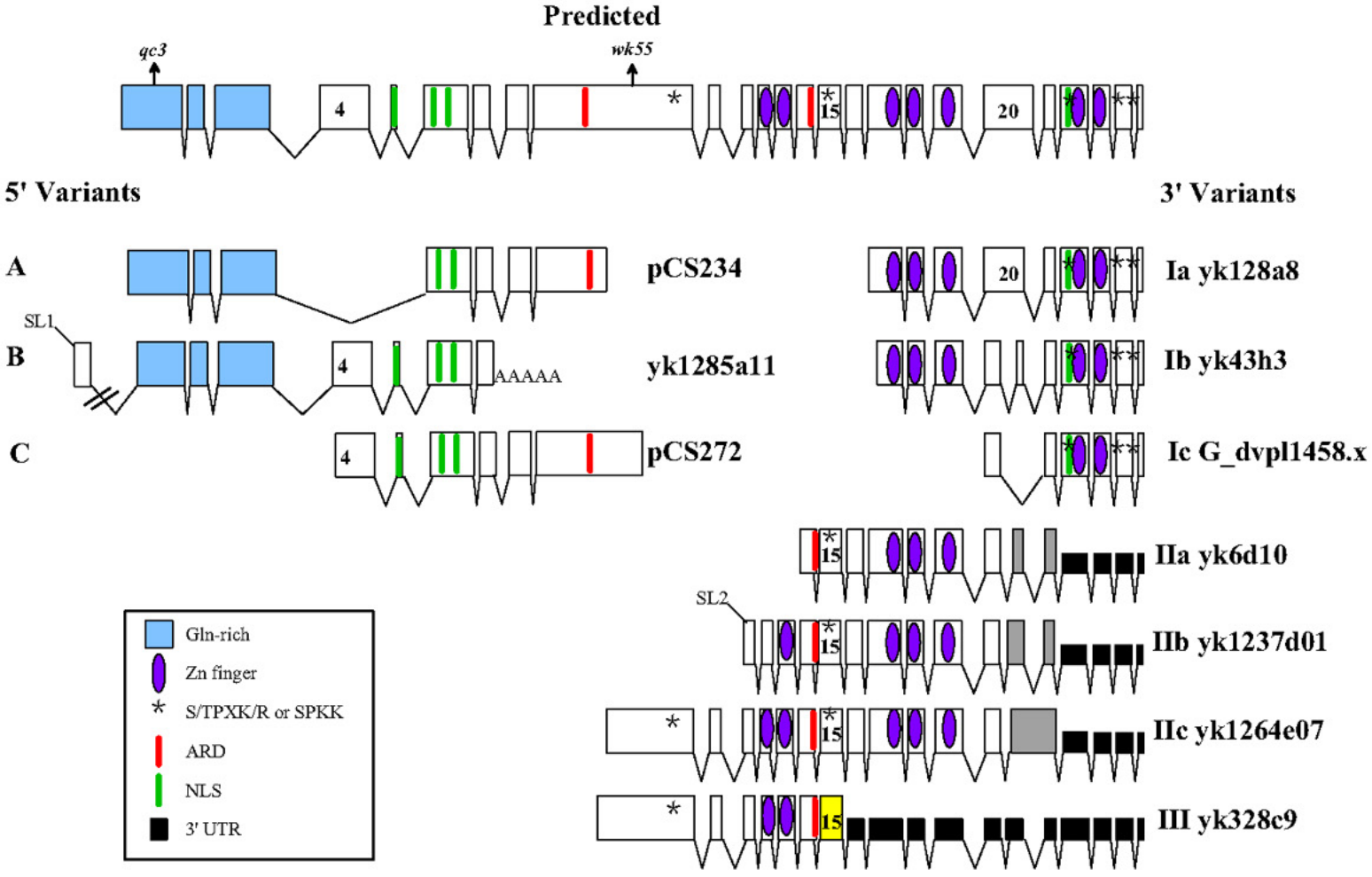
sma-9 displays complex alternative splicing

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 (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 (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

Fig. 3. Alternative splicing of *sma-9* transcripts (from Liang et al., 2003)

sma-9 cDNA clones reveal a complex alternative splicing at both 5'- and 3'- ends. At least three forms of N-terminus and seven of C-terminus 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. Blue indicates the Gln-rich domain; green bars are NLS; red bars are ARD; asterisks are S/TPXK/R or SPKK motifs; purple ovals are zinc finger motifs; gray boxes represent unique C-terminal sequences; yellow indicates an alternative exon 15 with two extra bases; smaller black boxes are 3' untranslated sequences. 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*, *yk1109f01*; Class Ib – *yk43h3*; Class Ic – *G-dvp11458.x* (Walhout et al., 2000); Class IIa – *yk6d10*, *yk864c1*, *yk1134e06*; Class IIb – *yk1237d01* that is SL2-spliced, *yk856b10*, *yk1057a6*, *yk1216e10*, *yk1103h10* and *yk127d10* that is also missing the intron between exon 21 and 22 (not shown); Class IIc – *yk1264e07*; Class III – *yk328c9*. Only the trans-spliced cDNA clones *yk1285a11* and *yk1237d01* appear to be full length.


Predicted
qc3
wk55
4
15
20
5' Variants
3' Variants
A
pCS234
Ia yk128a8
SL1
B
yk1285a11
Ib yk43h3
AAAAA
C
pCS272
Ic G_dvpl1458.x
IIa yk6d10
SL2
IIb yk1237d01
IIc yk1264e07
III yk328c9
Gln-rich
Zn finger
S/TPXK/R or SPKK
ARD
NLS
3' UTR

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 (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; 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 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. Because these cDNAs were rare, however, they might represent

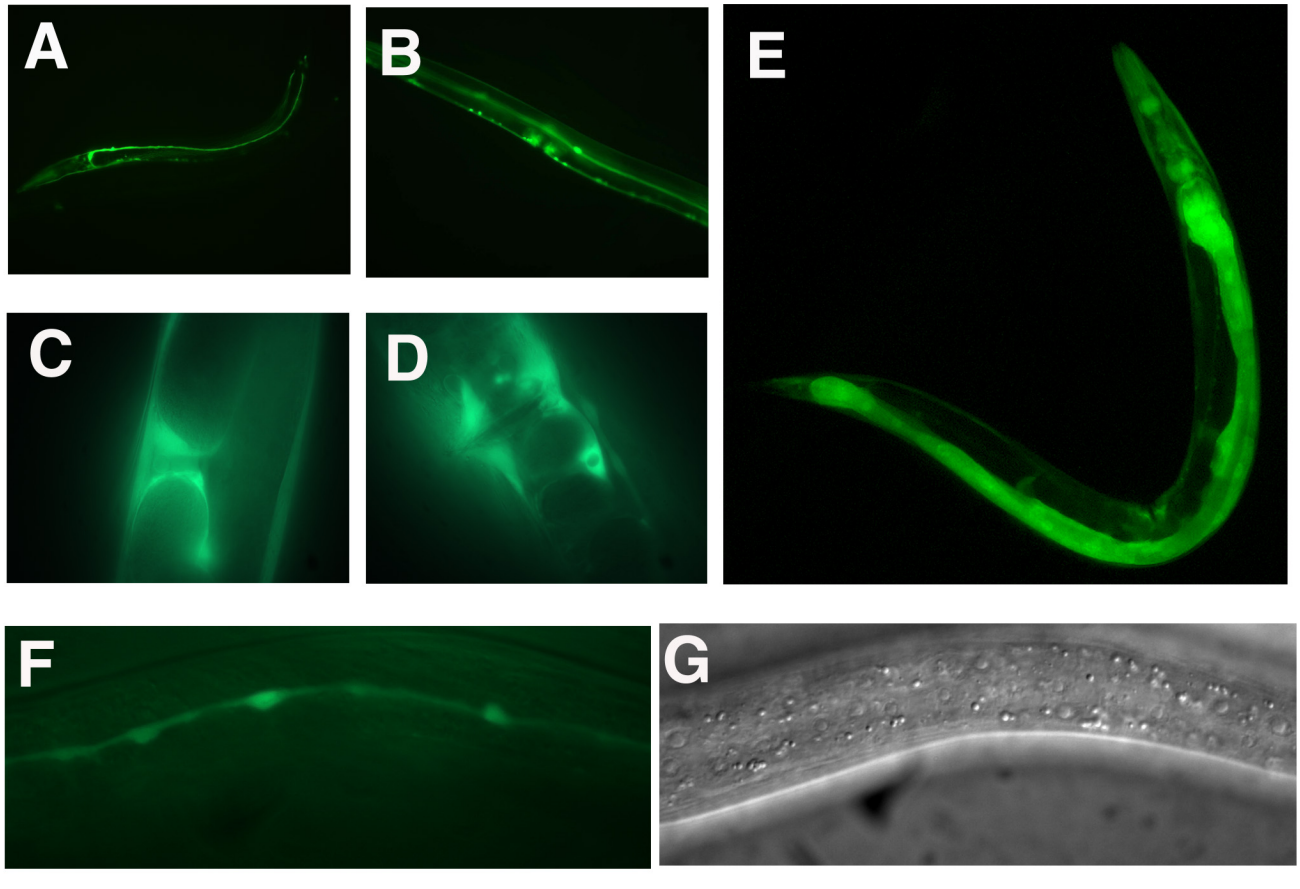
low abundance messages, messages with cell-type specific expression patterns, or spurious events with no functional significance. The region between exons 7 and 11 has some features of an intercistronic region, but is clearly not typical. Forty-eight nucleotides downstream of exon 7 is an imperfect match 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). The distance between the two transcripts, however, is 2.3kb, whereas a typical intercistronic region is between 100 and 120bp (Blumethal et al., 2002). The results of RNAi are suggestive 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 while 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 is 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 (Fig. 4). *sma-9* promoter-driven GFP was expressed in the excretory

Fig. 4. *sma-9* is widely expressed

sma-9 upstream sequences (H) were fused with a GFP reporter gene. (A-G) GFP expression in wild-type transgenic animals. The adult transgenic animals display a strong fluorescence in excretory canal (A), VNC (B), spermatheca (C), vulva (D), pharynx, and intestine (E), from L1 to adult. L2 animals show expression in the seam cells that disappears after L3 (F; Nomarski image in G). (H) shows different *sma-9* upstream regions used to construct GFP reporters.


A
B
E
C
D
F
G
H
8.0kb upstream
5.5kb upstream
2.8kb upstream
alternative
exon1
predicted
exon1
1.5kb upstream
4.0kb upstream

canal (Fig 4A), ventral nerve cord (VNC; Fig. 4B), spermatheca (Fig. 4C), vulva (Fig. 4D), pharynx and intestine (Fig. 4E), and seam cells (Fig. 4F). The expression appears from L1 to adult but not during embryonic development. Significantly, expression in the lateral seam coincided with the critical 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 L3 stages and not in L4 or the adult. 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 2.8kb, 5.5kb and 8.0kb constructs (upstream of predicted exon 1; Fig. 4H) show strong fluorescence in all detected tissues except the seam cells. The 1.5kb construct (-5500 to -4000 relative to predicted exon 1) displays expression in the seam cells, VNC, and excretory canal only. The 4.0kb construct (-8000 to -4000) generates the complete expression pattern, indicating the presence of redundant transcriptional elements. The 5.5kb and 8.0kb fragments are identical to the 1.5kb and 4.0kb fragments, respectively, except for the addition of sequences from -4000 to -1 that include the alternative exon 1 (Fig. 4H). 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 and -1. 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 5.5kb and 8.0kb constructs out of frame.

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 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). In 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 but is dispensable later.

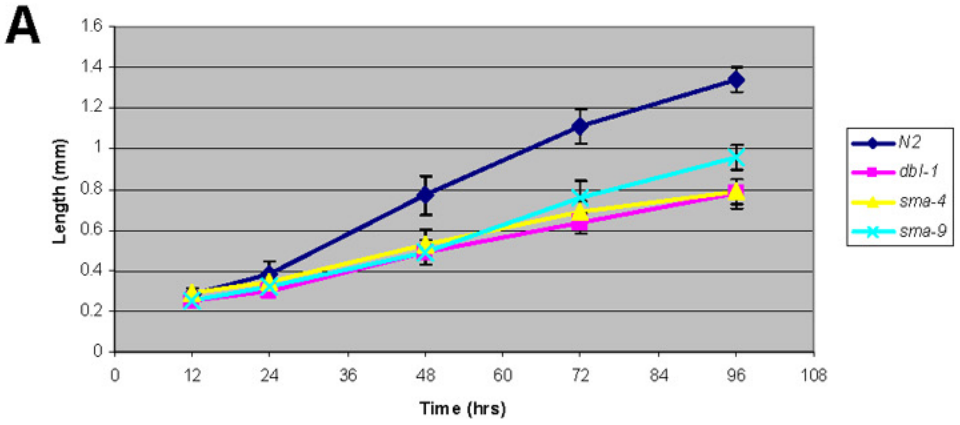
To better understand the relationship between *sma-9* and the DBL-1 pathway, we analyzed the effect of *sma-9*(*lf*) on *dbl-1* overexpression and *lon-1*(*lf*) phenotypes (Fig. 1). Overexpression of the ligand gene *dbl-1* from an integrated array in a wild-type background gives the animal a Lon phenotype (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;

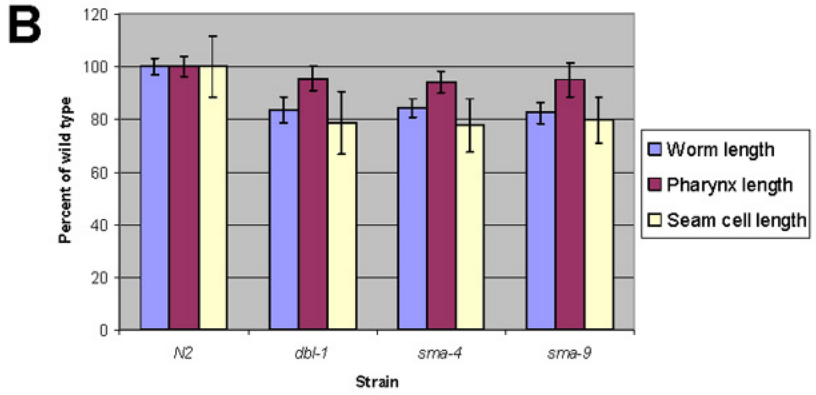
Fig. 5. *sma-9* functions in early stages to regulate body size (from Liang et al., 2003)

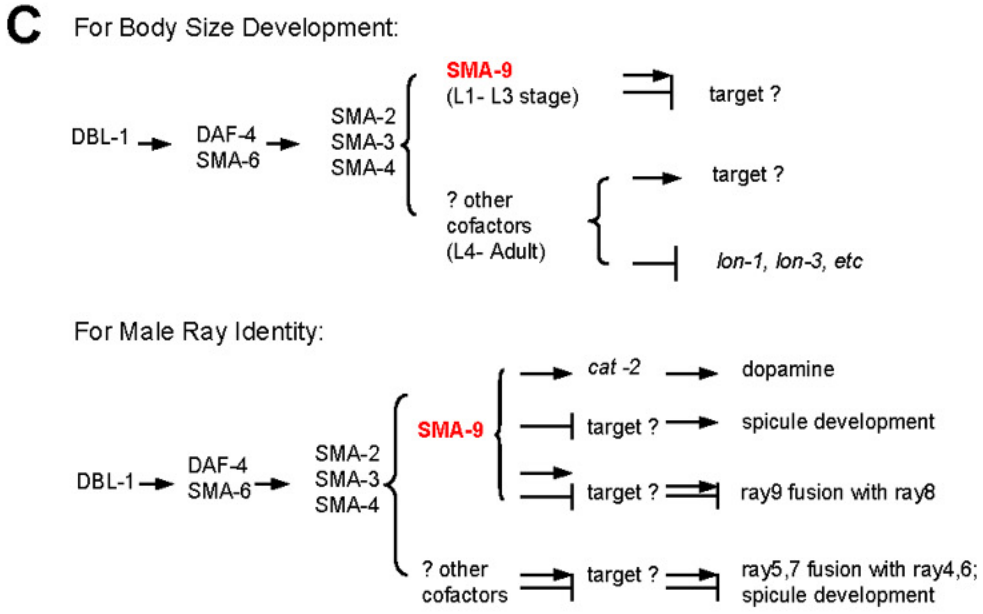
(A) Growth curve of *sma-9* (cross), *sma-4* (triangle), *dbl-1*(square) and N2 (diamond). In L1 larvae (12 hours), all animals have the same length. In L1 through L3 stages (48 hours), *sma-9, sma-4* and *dbl-1* show reduced growth rate compared with N2. After L3 *sma-9* mutants grow rapidly, while *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., 2000.

(B) At L3 stage, *sma-9* worm length, pharynx length and seam cell size is indistinguishable from that of *dbl-1* and *sma-4*. Data for N2 and *dbl-1* is from Wang et al., 2002.

(C) Model of *sma-9* function in 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 the 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 likely required.


A
Length (mm)
Time (hrs)
N2
dbl-1
sma-4
sma-9


B
Percent of wild type
Strain
Worm length
Pharynx length
Seam cell length


C
For Body Size Development:
DBL-1
DAF-4
SMA-6
SMA-2
SMA-3
SMA-4
SMA-9
(L1- L3 stage)
target ?
? other
cofactors
(L4- Adult)
target ?
lon-1, lon-3, etc
For Male Ray Identity:
DBL-1
DAF-4
SMA-6
SMA-2
SMA-3
SMA-4
SMA-9
cat -2
dopamine
target ?
spicule development
target ?
ray9 fusion with ray8
? other
cofactors
target ?
ray5,7 fusion with ray4,6;
spicule development

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 adulthood. Unlike *lon-1*; DBL-1 pathway mutants which are Lon, *lon-1;sma-9* doubles are neither Lon nor Sma but 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.

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 MBP-1/ PRDII-BF1, and mouse αA-CRYBP1. Based on the similarity of *sma-9(lf)* and DBL-1 pathway mutant phenotypes, their genetic interactions, and 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 L1 to L3, the *sma-9* mutant body size is indistinguishable from that of DBL-1 pathway mutants. Both *sma-9* and DBl-1 pathway component regulate male ray morphogenesis and spicules development. In the male rays both *sma-9* and DBl-1 signaling induce dopaminergic 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 (Liang et al., 2003). Furthermore, *sma-9(lf)* can suppress the

effects of *dbl-1* overexpression in both body size and male tail phenotypes (Fig. 1; Liang et al., 2003). These results 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 L1 to L3 stage, but is dispensable at later stages (Fig. 5). Since 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 critical 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 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 loss of other DBL-1 pathway components. *sma-9* mutants display ray 8-9

fusions at high frequency, but never ray 4-5 and 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 utilization of diverse

regulatory sequences 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 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, manuscript in preparation).

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. However, since *sma-9* mutants are viable, we have been able to analyze phenotypes throughout the course of development and make the assessment that the *sma-9* mutant phenotype is not weaker but rather 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 beginning to be appreciated.

CONTRIBUTIONS: The work was previously published (Liang J, Lints R, Foehr ML, Tokarz R, Yu L, Emmons SW, Liu J, Savage-Dunn C. (2003) The Caenorhabditis elegans schnurri homolog sma-9 mediates stage- and cell type-specific responses to DBL-1 BMP-related signaling. Development. 130(26):6453-64). Reprint was permitted by publisher. The genetic screens for *sma-9* alleles were performed by Dr. Cathy Savage-Dunn and Dr. Robyn Lints. 13 *sma-9* alleles's Mab phenotypes were characterized by our former lab member Rafal Tokarz. Meanwhile, he also measured the DBL-1 pathway mutant, *sma-1,* and *sma-9* L3 worm length and seam cell size, as well as their growth curves. Then *sma-9* was mapped to chromosome X at position +1.91- 2.2 by technician Ling Yu and former lab members. Dr. Jun Liu sequenced all the *yk* cDNA clones.

Chapter 2. *sma-9* acts as a transcriptional repressor in DBL-1/BMP Signaling pathway regulated body size development and male tail morphogenesis

ABSTRACT

In *C. elegans,* DBL-1/BMP pathway, a member of TGF-β superfamily, regulates body size development and male tail morphogenesis. Pathway component *sma-9* encodes a zinc finger transcription factor homologous to *Drosophila* Schnurri, which is required for Dpp/BMP signal. Our genetic studies using artificial SMA-9 give the first evidence that a transcriptional repressor can substitute for functional Shn-related protein *in vivo.* We show that the *sma-9* N-terminal Gln-rich region contributes to this transcriptional repressor activity by an *in vitro* transcriptional activity assay. Furthermore, among the total genes regulated by *sma-9* and *dbl-1* from our microarray analysis, there are more repressed genes than activated genes. Taken together, our results suggest that *sma-9* functions as a transcriptional repressor in DBL-1 pathway regulated body size development and male tail morphogenesis.

INTRODUCTION

Cell signaling pathways can elicit appropriate responses either by direct modification of cellular components, such as the cytoskeleton, or by the modulation of cell transcription profiles. Smad proteins are signal transducers for TGF-β (transforming growth factor-β) superfamily ligands that function by regulating gene transcription. Smad complexes activate or repress gene transcription depending on whether they associate with transcriptional activators or repressors. Since the DNA binding site for Smad

proteins show low complexity and affinity, transcriptional cofactors are required for target gene regulation (Shi et al., 1998; Massague, 1998; Derynck R and Zhang YE, 2003). We have identified a new component, *sma-9*, in a BMP/TGF-β pathway in *C. elegans.* SMA-9 protein products show motifs consistent with a role in transcriptional regulation, including a Gln-rich domain, seven zinc finger motifs, an acidic residue-rich domain and nuclear localization signals. Furthermore, SMA-9 is homologous to *Drosophila* Schnurri, which has been proposed to function as a transcription factor in Dpp signaling. In this study, we have addressed the molecular mechanisms of Shn protein function in BMP/TGF-β signaling in *C. elegans.*

The *C. elegans* DBL-1 pathway, a BMP-related signaling pathway, regulates body size development and male tail morphogenesis (Savage-Dunn, 2001). Loss of function of the DBL-1 pathway gives rise to small body size (Sma) and abnormal male tail (Mab) phenotypes. The major components in the pathway have been cloned: ligand *dbl-1*, type I receptor *sma-6*, type II receptor *daf-4*, R-Smads *sma-2* and *sma-3*, and Co-Smad *sma-4*. In addition, putative downstream components *sma-9, lon-1,* and *lon-3* have been cloned (Liang et al., 2003; Maduzia et al., 2002; Morita et al., 2002; Nystrom et al., 2002; Suzuki et al., 2002). We have shown that *sma-9* acts downstream of *dbl-1* in both body size and male tail patterning. Furthmore, SMA-9 is nuclearly localized, consistent with a function as a transcription factor. Based on phenotypic characterization, we have proposed that *sma-9* activity is required for a subset of *dbl-1*-mediated responses.

In the extensively studied Dpp/BMP pathway, Shn has been shown to physically bind to the Smad protein Mad (Dai et al., 2000) and repress expression of *brinker (brk)*- a default transcriptional repressor in the Dpp pathway (Marty et al., 2000; Muller et al.,

2003). *brk* had been considered to play a key role in pathway regulated gene transcription. The major function of Shn was proposed to be repression of *brk*. However, some evidence showed that Shn activates downstream target genes in a *brk*-independent manner (Torres-Vazquez et al., 2001). Furthermore, Pyrowolakis et al. reported that Dpp pathway regulated transcriptional repression was completely dependent on a short *cis*-acting silencer element (SE) identified originally from *brk* upstream sequences (Pyrowolakis et al., 2004). The SE has high affinity for Mad and Medea and subsequently recruits transcriptional repressor Shn. The authors found that many Dpp pathway target genes contain this sequence, which suggests a *brk*-independent transcriptional repressor activity of Shn. However, it is still unknown whether a functional transcriptional repressor can replace Shn during development. Since *Drosophila* Dpp pathway mutants are embryonic lethal, we took advantage of viable DBL-1 pathway mutants in *C. elegans* to address this issue. To date, no *brk* homolog has been found in the *C. elegans* genome.

To address the nature of the SMA-9 transcriptional activity, we have created SMA-9 fusions to known transcriptional activator and repressor domains: VP-16-SMA-9 and engrailed (En)-SMA-9 constructs were made, and introduced into worms. En-SMA-9 (a transcriptional repressor) rescued *sma-9(lf)* body size and male tail defects and caused a ligand overexpression phenotype in the male tail, suggesting that *sma-9* functions as a transcriptional repressor in DBL-1 pathway regulated body size and male tail morphogenesis. This is the first evidence that a heterologous transcriptional repressor can substitute for functional Shn protein *in vivo*. Meanwhile, we also showed that the N-terminal Gln-rich region of *sma-9* has transcriptional repressor activity. From our microarray data there are more genes repressed than activated by *sma-9*, which further

supports the hypothesis that *sma-9* acts as a transcriptional repressor. Additionally, among *dbl-1* regulated genes, more are repressed than are activated, suggesting that the DBL-1 pathway represses target genes in early larval development.

MATERIAL AND METHOD

Strains

C. elegans strains were cultured using standard methods and grown at 20°C (Brenner, 1974). In addition to strains generated in this work, the following strains were used: N2 (wild type), LG II: *sma-6(wk7);* LG III: *sma-2(e297), sma-3(wk30), sma-4(e729);* LG IV: *dbl-1(wk70);* LG V: *him-5(e1490)*; LG X: *sma-9(wk55)*. Since *sma-9(wk55)* has the strongest loss of function phenotype, most of our work was done on this strain.

Molecular cloning

4.0kb *sma-9* promoter region (Liang et al., 2003) was cloned in pBluscript SK+ at Not I and Sac II sites (pCS301); and then 3.6 kb *sma-9* C-terminal genomic sequence, which starts from the first zinc finger region to the stop codon, was cloned into pCS301 at Spel I and Kpn I sites (pCS303); finally VP-16 (370bp) or Engrailed (900bp) transcription activity domain (from vector TCF-vp16 and TCF-Engrailed (gifts from Gary Struhl)) was inserted into pCS303 at Not I and Spel I sites (pCS304 – VP-16-*sma-9;* pCS305-Engrailed- *sma-9*) respectively.

Primers used for pCS303: GGACTAGTATAATTGGCGGCCAT; and

GGGGTACCCTATAAGTCTTCTTCGCTTATTAATTTTTGTTCGTTCGCAGCCTT

(with c-myc tag);

Primers used for pCS304: GGGCGGCCGCATGGTACAAGATAAT; and GGACTAGTCCCACCGTACTCGTC;

Primers used for pCS305: GCGGCCGCATGGCCCTGGAGGATCGC; and GGACTAGTCAGAGCAGATTTCTCTGG;

sma-9 ARD II and cDNA clones: *yk1285a11* and pCS272 were cloned into yeast vector pSH 2-1, which give rise to pCS327, pCS326, and pCS330 respectively. Then those constructs were cotransfected with pLGΔ312S, pJK1621, and pCK30 (gifts from S. Kuchin) respectively into yeast strain MCY 829.

Primers used for cloning ARD II: AGGATCCTATTTCAATGTTTCGTC; and GGAATTCTACCATAGAAATCAACCA.

T27F2.4 :: *gfp* PCR product was generated as described (Hobert, 2002).

Primers used for 2.2kb T27F2.4 promoter regions: ACTGCAGTCTGGCACACGGCGCCAATG; GATGCAATTTTTTTGATTGT; GAAAAGTTCTTCTCCTTTACTCATGGATCCAGGAATGAAATGTGGACTGA.

Primers used for GFP ORF from vector pPD117.01 (gift from A. Fire): GGATCCTATCGATTCGCGGC; GTGCCACCTGACGTCGGCGC; AAGGCCTTTTGGGCCCAAG.

Transgenic animals

The plasmid DNA was microinjected into the gonadal syncytia of hermaphrodites with *rol-6* as a marker (Mello et al., 1991). 10ng/μl plasmid DNA or PCR product was injected into *him-5(e1490)* or *sma-9(wk55)* respectively. Arrays were integrated into chromosomes using γ-irradiation. Transgenic animals carrying pCS305 (*en-sma-9)* are:

CS331- *sma-9(wk55);him-5(e1490);qcIs25*; CS332- *sma-9(wk55);him-5(e1490);qcIs26*; CS333- *sma-9(wk55);him-5(e1490);qcIs27*; CS334: *sma-9(wk55);him-5(e1490);qcIs28;* CS370- *him-5(e1490);qcEx115;* carrying pCS304 (*VP-16-sma-9)* are: CS348- *sma-9(wk55);him-5(e1490);Ex116* and CS336- *him-5(e1490);qcIs30*; CS337- *him-5(e1490);qcIs31*; CS338- *him-5(e1490);qcIs32;* CS339- *him-5(e1490);qcIs33;* carrying pCS255 (1.5kb *sma-9::gfp)* is: CS301- *him-5(e1490);qcIs20;* carrying pCS251 (2.8kb *sma-9::gfp)* is: CS304- *him-5(e1490);qcIs23;* carrying pCS256 (4.0kb *sma-9::gfp)* is: CS303- *him-5(e1490);qcIs22*. Vectors pCS251, pCS255, and pCS256 are described in Liang et al (2003). CS332 and CS339 are the representative transformed animal strains, which give rise to the similar phenotype as others.

RT-PCR

Total RNA was isolated by Trizol (GibcoBRL Cat No. 15596-026) from wild type, *sma-9 (wk55),* the transformed animals strains CS332 and CS339. Briefly, worms were collected and washed with M9 buffer; then 10ml Trizol per ml packed worm was added; worms were frozen in liquid nitrogen and then thawed at 37 °C, repeat this 3-6 times; finally stored worms at -80°C until ready to precipitate. After thawing, the frozen worms/Trizol mixture were rapidly vortexed 30sec and then let stand on ice for 30sec, repeat this 6-7 times. Let stand at Room Temperature (RT) for 5min; 2ml chloroform per ml packed worm were added; shake by hand 15 sec, let stand at RT 2-3min; then centrifuge at 12,000g for 15min; collect the aqueous phase and add 5ml isopropanol, gently mix the solution and let the RNA precipitate at RT 10min; collect the RNA pellet

by centrifuge at 9,000g for 10min; wash the pellet with 10ml 75% ethanol; air dry the pellet and then dissolve in DEPC water; finally stored at -80°C.

RT-PCR was performed using SUPERSCRIPT™ One-Step RT-PCR (Invitrogen). Primers used for *VP-16-sma-9* are: GGGCGGCCGCATGGTACAAGATAAT; and AGGATCCTATTTCAATGTTTCGTC;

Primers used for *en-sma-9* are: GGGCGGCCGCATGGCCCTGGAGGATCGC; and AGGATCCTATTTCAATGTTTCGTC.

Yeast transcription activity assay

The assay was performed as described (Kuchin and Carlson, 2003).

Yeast two hybrid system assay

Yeast two hybrid system was from Life Technologies Inc (PROQUEST ™ Two-Hybrid System). *sma-9* cDNA clones: *yk328c9; yk128a8; yk43h3; yk127d10; yk228h6; yk6d10;* and *pCS272* were cloned into yeast vector pDBLeu or pPC86 respectively, as well as *sma-2, sma-3, sma-4* full length, MH1, and MH2 domains.

***sma-9* DNA microarray**

sma-9(wk55), dbl-1(wk70), and wild type N2 (as control) were grown on 150mm plates; eggs were collected by treating the adult animals with hypochlorite; incubated at 20°C; wait until worms developed to L2 (judge the stage of worms by examining the Nomarski image of the animal vulva); harvest the worm by M9 buffer. Total RNA was extracted by standard method using Trizol as described above.

Gene Chip *C. elegan* Genome Array were purchased from Affymetrix (Lot 3005451), which contains about 22,150 unique transcripts. At least duplicate experiments were performed on each strain. Hybridization and scanning were performed by Core Laboratory (GCL) at Memorial Sloan-Kettering Cancer Center. The software used for analyzing data were: GCOS (downloaded from Affimatrix Company) and the Macro program (developed by GCL- Memorial Sloan-Kettering Cancer Center).

dsRNAi

lin-49 constructs pCS308 was generated as following: cloning 800bp *lin-49* ORF region into pBluscript SK+ vectors at BamH I and Kpn I; then digest and insert into pPD129.36 at Sac I and Kpn I sites. T27F2.4 ORF sequence was cloned into pBluscript SK+ and RNAi feeding vector pPD129.36 (a gift from A. Fire) respectively at Sac I and Kpn I sites, which generated pCS359 and pCS360. dsRNA were made by vector pCS359 as described at Liang et al 2003, and then microinjected into hermaphrodite gonad. RNAi feeding method by using pCS308 and pCS360 was performed as described (Kamath and Ahringer, 2003).

Primers used for cloning T27F2.4 are: AAGAGCTC AGATGCTCTTCAACTGAA and CCGGTACC AGGAATGGTTGATTTGA.

Primers used for pCS308 are: GGGGATCC CCACCGTTGCTGCACTGTCC;

and GGGGTACC TTCCGCTTCCTGAAGGTGTC.

RESULTS

en-sma-9, but not *VP-16-sma-9,* rescues body size and male tail morphogenesis defects in *sma-9* mutants

DBL-1 pathway loss of function (lf) mutants have small body size in both hermaphrodite and male. In addition, the mutant male has abnormal male ray fusions and crumpled spicules (Savage et al., 1996). Wild type male tails have an acellular structure called the fan, which is formed from the cuticle matrix (Sulston et al., 1980). Within the fan, there are 18 sensory rays that can be grouped into two independent sets. Each set contains 9 rays distinguished by their specific position and shape (Fig1A). DBL-1 pathway (lf) mutant males have frequent ray fusions of rays 4-5, 6-7, and 8-9. When *dbl-1* is overexpressed in wild type background, the animal has a long body size and the male contains ray 3-4 fusions (Suzuki et al., 1999).

sma-9 encodes a transcription factor homologous to *Drosophila* Schnurri and functions as a downstream component of the DBL-1 pathway in *C. elegans* (Liang et al., 2003). Like other DBL-1 pathway components, *sma-9*(*lf*) mutant has a small body size (Sma) and abnormal male tail (Mab). Evidences support that *sma-9* regulates body size development at early larval stages and male tail morphogenesis of rays 8 and 9. To investigate whether *sma-9* acts as a transcriptional activator or repressor in the DBL-1 pathway, two artificial *sma-9* constructs were created: *sma-9* DNA binding domain linked with VP-16 or Engrailed (En) transcription activity domain, and with *sma-9* promoter region. Here we refer them as *VP-16-sma-9* or *en-sma-9*. Since VP-16 and En have been shown to be a strong transcriptional activator and repressor respectively (Tolkunova et al., 1998; Sze et al., 1997), VP-16-SMA-9 and En-SMA-9 would act as a

transcriptional activator and repressor respectively. Then, these constructs were microinjected into both *sma-9(lf)* and wild type background animals. Both the Sma and Mab phenotypes of the transformed animals were investigated.

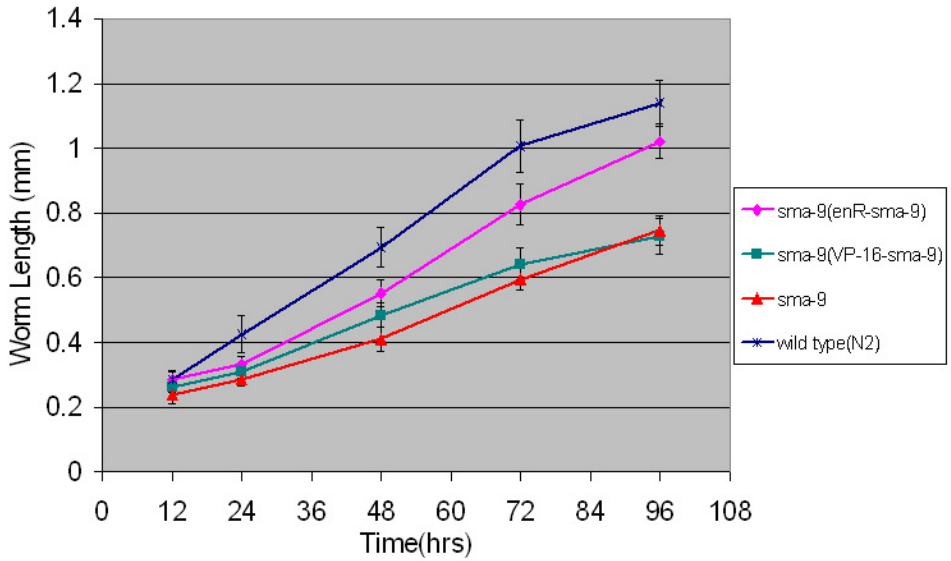
The *sma-9*(lf) mutants with *en-sma-9* constructs have longer worm length and more rapid growth rate than the mutant itself, although they still cannot reach the size of wild type animals (Fig 6A). The mail tail ray 8 and 9 fusions have been improved from 53% to 29%, and crumpled spicules were partially rescued from 46% to 8% (Table 2). Meanwhile a *dbl-1* overexpression phenotype, ray 3 and 4 fusion, was observed (Fig 6B). Furthermore, a higher penetrance of this ray 3-4 fusion occurred in animals with endogenous *sma-9*(+) (Table 2). These results suggest that *en-sma-9* can substitute for functional *sma-9*. Compared with *en-sma-9, VP-16-sma-9* transformed animals in both *sma-9* (lf) (Fig 6A) and wild type background do not change their body size, or male tail phenotypes (Table 2). Thus, *VP-16-sma-9* did not cause a dominant negative phenotype in wild type background. Since the *VP-16-sma-9* message was detected in the integrated animals, we believe that this is the real phenotype although we do not understand the reason. However, in *sma-9*(lf) background *VP-16-sma-9* message could not be detected, we are not completely sure whether this is the real phenotype or this results from the low expression level of the construct. Since *en-sma-9* but not *VP-16-sma-9* can partially rescue the body size and male tail morphogenesis defects in *sma-9* mutants, we propose that *sma-9* functions primarily as a transcriptional repressor in the DBL-1 pathway.

Fig.6. *sma-9* functions as a transcriptional repressor in the DBL-1 pathway regulated body size development and male tail morphogenesis.

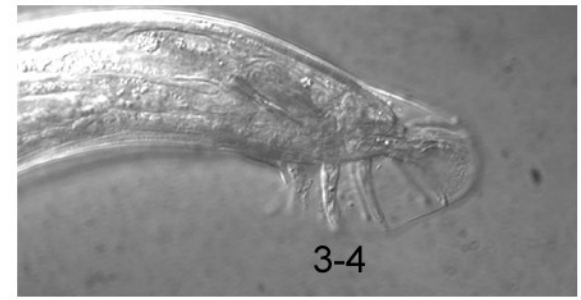
(A) Growth curve of *sma-9(wk55)* animals transformed with *En-sma-9* (diamond), *VP-16-sma-9* (square), and control (triangle). The *En-sma-9* transformed animals grow faster than *sma-9,* indicating that it partially rescues the body size. Meanwhile, *VP-16-sma-9* transformed animals grow the same as *sma-9*. The modified *sma-9* mRNA was detected by RT-PCR in transgenic strains CS332 (*en-sma-9*) (C-3) and CS339 (*VP-16-sma-9*) (D-3), but not in wild type (C-1; D-1) and *sma-9* (C-2; D-2) backgrounds.

(B) *En-sma-9* transformed animal gave rise to ray 3-4 fusions, which was reported in *dbl-1* overexpression strains (Suzuki et al., 1999). The fusion has a higher penetrance in *sma-9(+)* than in *sma-9(lf)* animals.

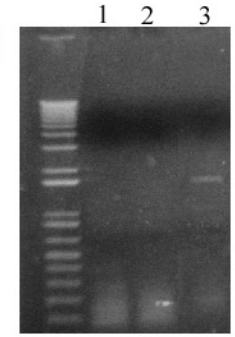
A


1.4
1.2
1
0.8
0.6
0.4
0.2
0
Worm Length (mm)
0
12
24
36
48
60
72
84
96
108
Time(hrs)
sma-9(enR-sma-9)
sma-9(VP-16-sma-9)
sma-9
wild type(N2)

B


3-4

C


1
2
3

D

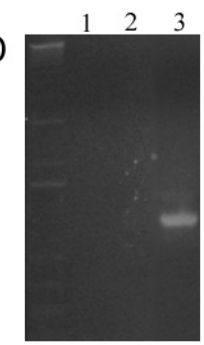

1
2
3

Table 2. En-SMA-9 rescued *sma-9(lf)* mutant male tail defects

Strains	Ray 3-4 fusion	Ray 8-9 fusion	N (sides)	Crumpled spicules	N (tails)	Expressed protein
wild type *him-5*	0%	16%	74	0%	30	None
mutant *sma-9*	0%	53%	101	46%	50	None
CS370 *him-5; qcEx115*	10%	23%	87	ND	ND	En-SMA-9
CS332 *sma-9; Is26*	Occasionally Observed	29%	87	8%	36	En-SMA-9
CS339 *him-5;qcIs33*	0%	8%	74	0%		VP-16-SMA-9
CS348 *sma-9; qcEx116*	0%	43%	81	50%	16	VP-16-SMA-9

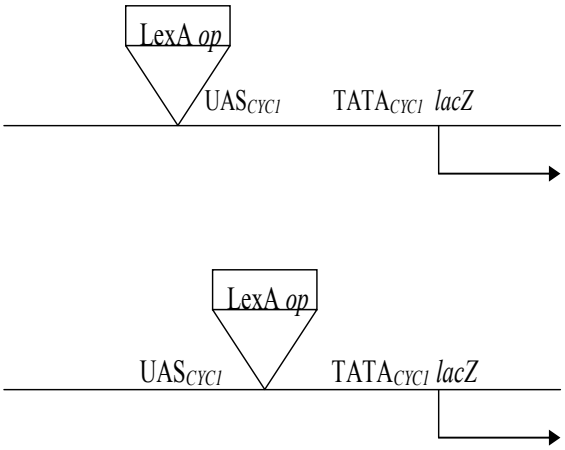
All strains in *sma-9* background are *sma-9(wk55)* alleles; in *him-5* background are *him-5(e1490)* alleles.

***sma-9* N-terminal region has transcriptional repressor activity**

To further test this hypothesis, we performed an *in vitro* transcription activity assay in a yeast system to investigate SMA-9 transcriptional activity (Fig 7; Table 3). Vector pLGΔ312S constitutively expresses *lacZ* gene whose expression level can be measured by β-Galactosidase (β-Gal) activity. Vector pJK1621 and pCK30 are derivatives that contain LexA operators. The tested gene fragments were fused with LexA DNA binding domain in yeast vector pSH 2-1. After cotransformation, the tested protein fragment should influence the *lacZ* expression via interaction between LexA DNA binding domain and its operator. Three fragments were used: Class B (*yk1285a11:* SL1 mediated full length 5' *sma-9* form including exons1-7), Class C (pCS272: 5' form including exon 4-9), and predicted exons 14 and 15 (Fig 3). Class B decreased β-Gal activity 1.3 fold, and Class C decreased activity 2.3 fold. Therefore, either the predicted exons 8 and 9 enhance Class C transcriptional repressor activity or the Gln rich region decreases the activity of Class B. These results show that *sma-9* N-terminal fragments have transcriptional repressor activity.

However, in predicted exons 14 and 15, the secondary acidic residue rich domain (ARD II), has maximum 1.6 fold increase of β-Gal exression. This indicates that ARD II contains transcriptional activator activity (Table 3). Since full-length *sma-9* cDNA clone *yk1237d01* excludes N-terminal sequences but contains ARD II and continuous DNA binding domain, it may have transcriptional activator activity *in vivo*.

Interestingly, our yeast two-hybrid assay supported these results. Originally, we performed yeast two-hybrid assay in order to investigate physical interaction between SMA-9 and Smad proteins: SMA-2, SMA-3, and SMA-4 (material and method).



Plasmid	No. of LexA operators
pJK1621	4
pLGΔ312S	0
pCK30	1
pLGΔ312S	0

Fig 7. Reporter plasmids used in yeast repression assays. There is no LexA operator in *CYC1-lacZ* reporter plasmid pLGΔ312S, which will constitutively express LacZ. Its derivatives pJK1621 and pCK30 have four or one LexA operators 5' or 3' to the UAS$_{CYC1}$.

Table 3. SMA-9 has transcriptional repressor activity

Fragment	β-Gal activity[1]		Fold Change[2]
	- LexA op	+ LexA op	
		5'to UAS	
pCS272	871.1 ± 61.1	378.9 ± 26.7	-2.3
ARD II	1512.7 ± 198.6	1826.3 ± 346.3	+1.2
		3'to UAS	
yk1285a11	1402.3 ± 307.7*	1081.5 ± 130.1	-1.3
ARD II	1512.7 ± 198.6	2494.8 ± 108.8	+1.6

[1.] β-Galactosidase activity was assayed in permeabilized cells and expressed in Miller units: Activity = (OD420 × 1,000)/(OD600 × vol assayed × time in min) for 3 to 4 independent clones. All standard errors were less than 13% (except * with 22%).

[2.] Fold change was calculated as the ratio of the values obtained for reporters with and without a LexA operator. – represent repression; + represent activation. The host strain was MCY 829.

However, no positive activities were detected from seven *sma-9* cDNAs, except that *yk6d10* clone was self-activating (data not shown). *yk6d10* does not contain N-terminal sequences but ARD II and continuous DNA binding domain. Consequently, *yk6d10* protein product has transcriptional activator activity, which constitutively activates reporter gene transcription no matter whether it physically binds to the other protein or not.

Since the other *sma-9* cDNA clones encode an N-terminal fragment that constitutively represses reporter gene transcription, then transcriptional activation may be blocked even if protein-protein interactions occur. Thus, both our *in vitro* transcription activity assay and yeast two-hybrid assay support that the *sma-9* N-terminal region contains transcription repressor activity.

sma-9 and _lin-49_ function synergistically in male ray morphogenesis in DBL-1 pathway independent manner

Drosophila Shn represses expression of *brk,* a default transcriptional repressor, in Dpp pathway through a silencer element (SE). Furthermore, mutation of *brk* suppresses *shn* mutant phenotypes. In the worm, the *lin-49* promoter region has a relatively similar SE sequence. Bromodomain protein LIN-49 is a transcription factor related to trithorax group (trx-G) proteins in *Drosophila.* Strong alleles are embryonic lethal, but weak alleles survive producing males with defects in the tail fan and bursa especially highly frequent fusion of rays 3 and 4 (Chamberlin and Thomas, 2000). Noticing that the *dbl-1* overexpression animals contain ray 3 and 4 fusions at a low penetrance, we asked

whether *lin-49* might be a downstream target gene or functions as *brk* in the DBL-1 pathway.

We performed *lin-49* dsRNAi by feeding in wild type and Sma mutants. The results show that *lin-49* RNAi itself does not cause any defect in wild type, probably due to the low efficiency of the feeding RNAi method. Meanwhile, *lin-49* RNAi did not affect body size of Sma mutants *sma-2, sma-3,* and *sma-9* (data not shown). However, investigation of *lin-49* dsRNAi in *sma-2* and *sma-3* and *sma-9* (lf) mutant males show that their male tails have much more severe ray fusion phenotypes (Table 4). Most rays were misshaped in addition to the original mutants phenotypes, for example, ray 1-2, 3-4, 1-2-3, 3-4-5, and even occasionally ray 1-2-3-4 fusions (Table 4). The results indicate that *lin-49* is not a repressed downstream target of *sma-9,* since knockdown of *lin-49* did not block *sma-9(lf)* defects.

Meanwhile, *lin-49::gfp* reporter expression levels in wild type and *sma-9 (lf)* were the same, suggesting that *sma-9* does not regulate *lin-49* transcription (data not shown). However, it suggests that the DBL-1 pathway and *lin-49* have synergistic effects on male ray morphogenesis, in particular rays 1, 2, 3, 4, and 5. The effects of *lin-49* RNAi on body size and male ray fusion patterns in all three Sma mutants, suggests that *lin-49* is not *C. elegans* functional *brk* and does not genetically function in but with the DBL-1 pathway for mail tail morphogenesis.

Table 4 *lin-49* RNAi Mab phenotypes in DBL-1 pathway (lf) background

Mutant	Fusion of Rays 4-5	Fusion of Rays 6-7	Fusion of Rays 8-9	Other Fusions	N (sides)
him-5(e1490)	0%	0%	16%	NA	74
single mutants *sma-3(wk30)*	27%	42%	15%	NA	97
sma-2(e297)	5%	78%	40%	NA	76
sma-9 (wk55)	0%	0%	53%	NA	101
***lin-49* RNAi in** *him-5(e1490)*	0%	0%	22%	NA	95
sma-3(wk30)	28%	37%	30%	1-2: 2.5% 1-2-3: 1% 2-3: 1% 3-4: 2.5%	79
sma-2(e297)	10%	36%	26%	1-2-3: 1% 3-4: 10%	86
sma-9 (wk55)	0%	0%	48%	1-2: 7% 1-2-3: 3% 3-4: 19% 3-4-5: 10%	176

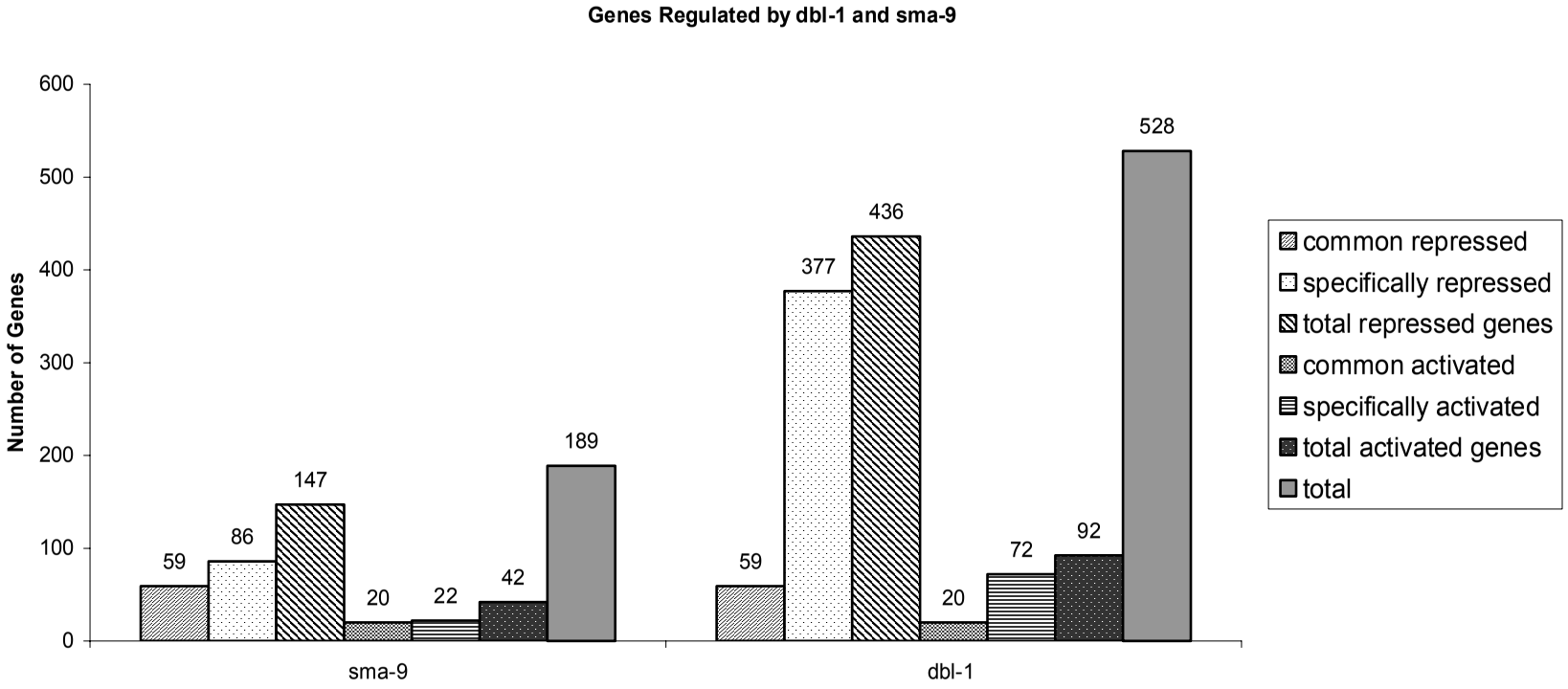
More genes were repressed than those were activated in DNA microrray profile

To identify potential *sma-9* transcriptional target genes in a more unbiased manner, we determined patterns of gene expression using a DNA microarray. Since *sma-9* functions in the early larval stage in body size regulation (Liang et al., 2003), L2 animals were used to perform DNA microarray analysis. Significant genes were selected with greater than 2 fold changes ($P< 0.001$) of expression level compared with wild type (Table 6, 7, 8). Because the animals used are loss of function mutants, genes upregulated in the mutant background are repressed by the specific gene's activity in the wild type background. Vice versa, genes downregulated in the mutant background are activated by the specific genes. The genes identified will include both direct and indirect target genes.

The DNA microarray data suggests that *sma-9* regulates a subset of DBL-1 pathway target genes. Total genes altered in *dbl-1* background are 528, in *sma-9* background are 189 (Fig 8). 79 genes were shared by both. Thus, *sma-9* only regulated a subset of target genes in the DBL-1 pathway. This is consistent with its specific function as we described previously that *sma-9* functions at early larval stages for body size development and regulates specifically rays 8 and 9 morphogenesis. We propose that in order to fulfill the complete function of the DBL-1 pathway, the Smad complex needs *sma-9* and other transcriptional cofactors.

Among the 189 genes directly or indirectly regulated by *sma-9,* 147 genes are repressed and 42 genes are activated. The results support our hypothesis that *sma-9* functions as transcriptional repressor. Surprisingly, among the 528 genes regulated by *dbl-1,* 436 genes are repressed but only 92 genes are activated. It suggests that at least during early developmental stages the DBL-1 pathway mostly represses target gene

Fig 8. Number of genes identified from DNA microarray analysis regulated by *dbl-1* and *sma-9*. More genes were repressed than were activated. *dbl-1* regulates more genes than *sma-9* does.


Genes Regulated by dbl-1 and sma-9
Number of Genes
600
500
400
300
200
100
0
59
86
147
20
22
42
189
59
377
436
20
72
92
528
sma-9
dbl-1
common repressed
specifically repressed
total repressed genes
common activated
specifically activated
total activated genes
total

expression. How does this happen? One possibility is that the DBL-1 pathway stimulates expression of a transcriptional repressor such as *sma-9*, which represses target gene transcription. To investigate whether *sma-9* is regulated by the pathway, different *sma-9* transcriptional reporter genes (material and method) were integrated into a wild type background, and then were crossed into *dbl-1, sma-6, sma-2, sma-3,* and *sma-4* backgrounds. However, none of them has significant expression change compared with wild type background (data not shown). Meanwhile, *sma-9* is not regulated by *dbl-1* from our microarray analysis. Thus, we exclude the possibility that the DBL-1 pathway represses target gene transcription by regulating *sma-9* expression level. Taken together, we propose that *sma-9* regulates a subset of DBL-1 pathway genes and functions as a transcriptional repressor in the early developmental stages.

Genes regulated by *sma-9* and *dbl-1* from microarray analysis

Previously identified genes

The gene K01A2.3, Y105C5A.13, and R09H10.5 have been reported to be regulated by the DBL-1 pathway through hybridization of cDNA arrays (Mochii M et al 1999) (Table 5). The authors showed that all three genes were downregulated in loss of function of *dbl-1* and *sma-2* mutants compared with wild type. All three genes were expressed in the intestine where many pathway components are located. In our array, all were downregulated in *dbl-1* background. From a high-throughput genome-wide RNAi analysis of gene function (Maeda I et al 2001), loss of Y54G9A.5 function gives rise to a

Table 5 Genes identified by Microarray profile in DBL-1 pathway

Gene Name	Identification	Average fold change	Background	Reference
Target genes				
K01A2.3	protein of unknown function	-3.28	*dbl-1*	Mochii et al., 1999
Y105C5A.13	protein of unknown function	-5.55	*dbl-1*	Mochii et al., 1999
R09H10.5	contains EGF domains	-2.38	*dbl-1*	Mochii et al., 1999
Y54G9A.5	negative regulation of body size	6.93	*dbl-1*	Maeda et al., 2001
egl-46	a TFIIA-like zinc finger transcription factor	-2.07	*dbl-1*	Maeda et al., 2001
Collagens				
col-88	cuticular collagen (type IV and type XIII)	18.93	*dbl-1*	
col-120	cuticular collagen (type IV and type XIII)	11.85	*dbl-1*	
col-175	cuticular collagen (type IV and type XIII)	9.09	*dbl-1*	
col-176	cuticular collagen (type IV and type XIII)	8.31	*dbl-1*	
Y77E11A.12	cuticular collagen (type IV and type XIII)	5.70	*dbl-1*	
bli-1	cuticular collagen (type IV and type XIII)	4.80	*dbl-1*	

col-37	cuticular collagen (type IV and type XIII)	5.73	*dbl-1*	
col-141	cuticular collagen (type IV and type XIII)	3.19	*dbl-1*	
col-142	cuticular collagen (type IV and type XIII)	2.65	*dbl-1*	
rol-6	cuticular collagen (type IV and type XIII)	-2.15	*dbl-1*	
col-41	cuticular collagen (type IV and type XIII)	-2.64	*dbl-1*	
col-80	cuticular collagen (type IV and type XIII)	2.23	*sma-9*	
Cell cycle				
C04C3.4	G2/Mitotic-specific cyclin A	2.75	*dbl-1*	
T07F10.5	G2/Mitotic-specific cyclin A	2.64	*dbl-1*	
F08F1.9	G2/Mitotic-specific cyclin A	2.57	*dbl-1*	
mcm-7	DNA replication licensing factor	-2.07	*dbl-1*	
cdk-1	cyclin-dependent kinase family	-2.07	*dbl-1*	
Innate immunity				
T07D10.4	C-type lectin	-2.95	*dbl-1*	Murphy et al., 2003
B0218.8	C-type lectin	-2.79	*dbl-1* and *sma-9*	Murphy et al., 2003
F35C5.9	C-type lectin	-3.85	*dbl-1* and	Mochii et al., 1999

			sma-9	
pgp-1	multidrug/pheromone exporter, ABC superfamily	-2.57	*dbl-1*	
pgp-6	multidrug/pheromone exporter, ABC superfamily	-2.89	*dbl-1*	
nlp-29	neuropeptide-Like Protein	3.38	*sma-9*	Mallo et al., 2002
lys-10	lysozyme	-4.27	*sma-9*	
spp-12	saoosin-like Protein family	354.95	*dbl-1*	Murphy et al., 2003
cnc-2	caenacin (Caenorhabditis bacteriocin)	6.50	*sma-9*	Mallo et al., 2002
cnc-4	caenacin (Caenorhabditis bacteriocin)	3.02	*dbl-1*	Couillault et al., 2004
Chemoreceptors				
C05E7.4	7TM chemoreceptor	2.07	*dbl-1*	
F01E11.5	7TM chemoreceptor	-5.36	*sma-9*	
F19B2.3	7TM chemoreceptor	5.04	*dbl-1* and *sma-9*	
K06C4.9	7TM chemoreceptor	2.07	*dbl-1*	
K08D10.10	7TM chemoreceptor	-2.52	*dbl-1*	

srd-27	7TM chemoreceptor	2.65	*sma-9*
srg-13	7TM chemoreceptor	3.38	*dbl-1*
srg-28	7TM chemoreceptor	3.04	*dbl-1*
srh-11	7TM chemoreceptor	2.07	*dbl-1*
srh-134	7TM chemoreceptor	3.19	*dbl-1*
srh-247	7TM chemoreceptor	3.19	*dbl-1*
srh-47	7TM chemoreceptor	2.67	*dbl-1*
srv-21	7TM chemoreceptor	3.02	*dbl-1*
srw-44	7TM chemoreceptor	2.32	*dbl-1*
srw-44	7TM chemoreceptor	2.38	*sma-9*
srw-60	7TM chemoreceptor	2.67	*dbl-1*
srx-14	7TM chemoreceptor	2.67	*dbl-1*
srx-77	7TM chemoreceptor	2.83	*dbl-1*
srz-85	7TM chemoreceptor	5.94	*dbl-1*
str-112	7TM chemoreceptor	2.38	*dbl-1*
str-244	7TM chemoreceptor	2.38	*dbl-1*

Y116A8C.41	7TM chemoreceptor	4.16	*dbl-1*	
R07C12.1	7-transmembrane receptor	2.52	*sma-9*	
sra-33	chemosensory perception	2.07	*dbl-1*	
srb-1	sra family integral membrane protein	4.83	*dbl-1*	
srv-8	weak similarity to 7-transmembrane receptor	2.86	*dbl-1*	
srx-16	7-transmembrane receptor	3.87	*dbl-1*	
T27C5.10	chemoreceptor	3.38	*dbl-1*	
gpa-8	G-protein alpha subunit	2.22	*dbl-1*	
Dauer and Aging genes				
vit-3	yolk protein	4.33	*dbl-1*	
vit-5	yolk protein	15.51	*dbl-1*	Murphy et al., 2003
vit-6	yolk protein	5.29	*dbl-1*	
fat-6	acyl-CoA desaturase	-3.25	*dbl-1*	Murphy et al., 2003
fat-7	fatty acid desaturase	4.31	*dbl-1* and *sma-9*	Murphy et al., 2003
dod-3	downstream of DAF-16	-2.87	*dbl-1*	Murphy et al., 2003

dod-21	downstream of DAF-16	3.30	*dbl-1*	Murphy et al., 2003
ins-7	insuline-like peptide	4.23	*dbl-1* and *sma-9*	Murphy et al., 2003; Liu et al., 2004
ins-4	Insulin-like peptides	2.93	*dbl-1*	Liu et al., 2004
daf-9	cytochrome P450	3.26	*dbl-1*	Liu et al., 2004
sip-1	heat shock hsp20 proteins	13.36	*dbl-1*	Murphy et al., 2003; Hsu et al., 2003
sod-3	superoxide dismutase	3.10	*dbl-1*	Murphy et al., 2003
hedgehog/patched				
wrt-4	warthog	-2.95	*dbl-1* and *sma-9*	Liu et al., 2004
wrt-9	warthog	-2.14	*dbl-1*	
grl-23	ground-like	25.85	*dbl-1*	
grl-25	ground-like	97.18	*dbl-1*	
ptc-3	patched related family	-2.38	*dbl-1*	Liu et al., 2004
ptr-22	patched related family	3.22	*dbl-1*	
major sperm protein				
msp-3	major sperm protein	14.96	*dbl-1*	

msp-10	major sperm protein	26.91	*dbl-1*
msp-31	major sperm protein	22.95	*dbl-1*
msp-33	major sperm protein	29.69	*dbl-1*
msp-36	major sperm protein	28.85	*dbl-1*
msp-38	major sperm protein	27.08	*dbl-1*
msp-50	major sperm protein	21.90	*dbl-1*
msp-74	major sperm protein	8.90	*dbl-1*
msp-76	major sperm protein	39.95	*dbl-1*
msp-77	major sperm protein	31.20	*dbl-1*
msp-78	major sperm protein	20.36	*dbl-1*
msp-79	major sperm protein	31.22	*dbl-1*
msp-81	major sperm protein	23.18	*dbl-1*
msp-152	major sperm protein	74.66	*dbl-1*
ZK1248.4	major sperm protein domain	10.82	*dbl-1*
F44D12.7	major sperm protein domain	10.32	*dbl-1*
C35D10.11	major sperm protein domain	9.52	*dbl-1*

cav-1	caveolin protein family	-2.98	*dbl-1* and *sma-9*
Germline			
B0280.5	a direct target of GLD-1	3.70	*dbl-1* and *sma-9*
cej-1	a direct target of GLD-1; cell junction protein;	3.88	*dbl-1*
daz-1	human DAZ (Deleted in Azoospermia) homolog;	2.52	*dbl-1*
ego-1	enhancer of Glp-1; RNA-directed RNA polymerase	3.06	*dbl-1*
fbf-2	RNA-binding protein	2.75	*dbl-1*
fog-3	feminization Of Germline	2.00	*sma-9*
gld-1	defective in Germ Line Development	2.76	*dbl-1* and *sma-9*
gld-3	defective in Germ Line Development	19.59	*dbl-1*
glh-1	germ-line helicase; DEAD-box RNA helicase	2.15	*dbl-1* and *sma-9*
glh-2	germ-line helicase	2.00	*sma-9*
him-3	high Incidence of Males	2.62	*dbl-1* and

			sma-9
hop-1	homolog of presenilin	2.30	*dbl-1* and *sma-9*
iff-1	Initiation Factor 5 (eIF-5A) homologs	2.43	*dbl-1* and *sma-9*
rec-8	a meiosis-specific cohesin complex subunit	2.22	*sma-9*
syp-3	coiled-coil protein functions Synapsis in meiosis	2.38	*dbl-1*

long (Lon) phenotype. Thus, Y54G9A.5 negatively regulates animal body size growth. In our analysis, it was upregulated. Consequently, our results suggest that these four genes might be potential targets for DBL-1 pathway regulated body size development.

Collagens

Cuticle collagens are synthesized by the hypodermis and secreted on the surface of the worm body as an exoskeleton. In the *C. elegans* genome, about 175 genes encode small collagen-like polypeptides. Mutation of these genes causes defects in body morphogenesis, e.g. dumpy, roller, blister, and embryonic lethality (Myllyharju and Kivirikko, 2004). The collagen gene *lon-3* negatively regulates body size development as a possible downstream component in the DBL-1 pathway (Suzuki et al., 2002). Loss of *lon-3* function gives rise to a Lon phenotype. Collagen genes *rol-6* and *sqt-1* genetically interact with one another for animal morphology (Kramer et al., 1990). It has been reported that both are required for *dbl-1* regulated body length and genetically interact with *lon-3,* and these interactions are reciprocal (Nystrom et al., 2002). Loss of *rol-6* or *sqt-1* function alone hardly changes the animal body size. However, either *rol-6* or *sqt-1* mutants can suppress *lon-3(lf)* and *dbl-1(*over expression) phenotype. In our data *rol-6* is downregulated in *dbl-1* mutants, consistent with these findings. Totally 11 collagen genes were regulated by *dbl-1*, and only one by *sma-9* (Table 5). The majority of collagen genes are upregulated in Sma mutant background. Thus, the results further support that collagen genes play an important role in DBL-1 pathway regulated body size development. However, *sma-9* might not play a major role in regulating collagen gene expression.

Cell cycle control genes

cdk-1 and *mcm-7,* cell cycle control genes, were downregulated in *dbl-1* mutant (Table 5). *cdk-1* is a homolog of yeast *cdc2* (also know as *cdk-1*) that controls G2/M phase transition of the cell cycle. *mcm-7* is homologous to minichromosome maintenance proteins (MCM) that play a critical role in S phase genome stability. MCM are DNA replication licensing factors in eukaryote cells. Studies from yeast, *Xenopus,* and mammalian cells demonstrated that MCM are targets of the S phase checkpoint and are essential for DNA replication initiation and elongation (Bailis and Forsburg, 2004). Both *cdk-1* and *mcm-7* loss of function by RNAi leads to embryonic lethality, showing that they are essential for *C. elegans* development.

The DBL-1 pathway (lf) mutants grow more slowly than wild type, which results in a small body size. It has been reported that hyp7 (hypodermal syncytium) is a critical tissue for body size development, due to the facts that expression of *sma-3, sma-6,* and *lon-1* in hyp7 alone is sufficient to rescue the respective mutant body size defects (Wang et al., 2002; Yoshida et al., 2001; Maduzia et al., 2002). Hyp7 is a single multinucleate hypodermal cell that extends over most of the animal body. In L1 larvae, 23 diploid nuclei exist in hyp7. Afterwards, daughter cells of P blast cells and the seam cells undergo endoreplication and become tetroploid, and then fuse with hyp-7. Totally, 110 cells fuse together and form a single syncytium. In DBL-1 pathway mutants the cell number does not change, indicating that a defect in cell division is not the cause of small body size (Suzuki et al., 1999, Flemming et al., 200). However, several lines of evidence showed that the DNA content in polyploid hypodermal nuclei due to endoreplication was reduced in the pathway mutants (Flemming et al., 2000; Morita et al., 2002). Perhaps

downregulation of *mcm-7* blocks endoreplication in intestine or hypodermal cells. Alternatively, the low expression level of both *cdk-1* and *mcm-7* could still be able to allow cell cycle progression, however it might lengthen the whole process. Careful investigation of DBL-1 pathway (lf) mutants supports this idea. Compared with wild type animals, the development of Sma mutants is delayed (data not shown). It takes them a longer time than wild type to develop to the same stage.

Genes related to innate immunity

DBL-1 pathway mutants have been shown to be more susceptible to bacterial infection than wild type animals (Kurz and Tan, 2004). In our microarray analysis, we identified genes that affect innate immunity, similar to previous studies (Mallo et al., 2002; Nicholas and Hodgkin, 2004) (Table 5). Type-C lectins are cell surface proteins capable of binding carbohydrate (Loris, 2002). In the *C.elegans* genome, about 135 genes encode a type-C lectin domain. Worms would use these numerous type-C lectins to recognize different pathogens. In our array, many type-C lectin genes were regulated by *dbl-1* and *sma-9* including the genes F35C5.9, B0218.8, and T07D10.4 (Table 5) identified from previous studies (Murphy et al., 2003; Mochii et al., 1999). Compared with previous reports (Mallo et al., 2002; Couillault et al., 2004), *nlp-29, cnc-4, cnc-2, lys-10,* and *spp-12* expression changed in the opposite direction. However, since *nlp-29, nlp-31, cnc-4,* and *cnc-2* were induced specifically upon bacterial *S. marcescens* or fungal *D. coniospora* infection of adult animals (Couillault et al., 2004), they represent a pathogen specific response in *C. elegans,* so perhaps these genes are not the primary triggers of the innate immunity response. In addition, two members of the PGP (P-glycoprotein) subclass of ATP-binding cassette (ABC) transporter family have been

identified in our array: *pgp-1* and *pgp-6.* ABC transporter actively translocates various molecules across the cell membrane. In humans, its functions include protecting cells from toxins and resistance to drugs e.g. antibiotic, anticancer drugs etc (Higgins and Linton, 2004). In *C. elegans, pgp-1(lf)* mutant, together with *mrp-1* or *pgp-3,* was hypersensitive to heavy metals or bacteria, indicating that *pgp-1* plays a role in a defense response. The function of *pgp-6* is still not characterized. In our analysis, both *pgp-1* and *pgp-6* were downregulated in the *dbl-1(lf)* mutant, which might contribute to the mutant's high susceptibility to bacterial infection.

Chemoreceptors

In the worm defense system, the animal might first sense the pathogen from the environment through sensory neurons (Millet and Ewbank, 2004). G-protein coupled receptors (GPCR) mediate chemosensation both in vertebrates and invertebrates (Troemel et al., 1995, Xu et al., 2004; Oka et al., 2004). The *C. elegans* genome encodes about 1000 GPCR members that mediate a large variety of chemical sensations. We found many chemoreceptors in our microarray profile. The majority of them were upregulated, although some were downregulated (Tables 5 & 8). Thus, it indicates that Sma mutants could be more sensitive to odor or taste cues than wild type. This might explain the natural phenomenon that Sma mutants can not grow well in a liquid culture compared with wild type animals.

Dauer and aging genes

We see altered transcription of dauer and aging related genes, also identified in previous studies (Murphy et al., 2003; Liu et al., 2004; Table 5). In our array, genes extending lifespan were downregulated including *fat-7, fat-6* and *dod-3;* genes shortening

lifespan were upregulated including *ins-7, vit-5,* and *dod-21*. In addition to *vit-5,* two other yolk protein genes *vit-3* and *vit-6,* were also upregulated in the mutant background. It would be interesting to test whether they play a role in longevity too. However, the genes *sod-3* and *sip-1* changed in the opposite direction from previous reports (Murphy et al., 2003; Hsu et al., 2003). One reason might be that we chose L2 animals rather than adults; so a full comparison may not be possible. Many genes regulating longevity also control dauer formation. Our analysis also identified genes that function in the dauer pathway. For example, *daf-9* and *ins-4* were upregulated, which suppress dauer formation as *ins-7* does.

The *dbl-1* mutant has been shown to live shorter than wild type on standard worm plates using *E. coli* OP50 as a food source (Mallo et al., 2002), which is a pathogen for the worm. The genes that we found affecting aging could also contribute to a shortened lifespan. Together, these results support that the loss of function of the DBL-1 pathway leads to a shorter lifespan in a diet of OP50.

hedgehog

The Hedgehog (Hh) family and its receptor (Patched and Smoothened) play an essential role in *Drosophila* and vertebrate development (Tabata and Takei, 2004). The genetic interaction between Hh and TGF-β has been extensively studied in other organisms, particularly in *Drosophila* (Tabata and Takei, 2004; Roussa and Krieglstein, 2004; Silver and Rebay, 2005). However, how the two signaling pathways interact with each other in *C. elegans* is poorly understood. Worm Hh homologs contain highly conserved C- terminus (hog), but various N-termini with a low similarity. They are classified into different groups based on their N-terminal sequence as: *warthog(wrt),*

groundhog(grd), groundlike(grl), and *hh* (Aspock et al., 1999). There are two families of Hh receptor in the *C. elegans* genome: patched (*ptc*) and patched-related (*ptr*). Our data showed three hedgehog gene families: *wrt-4, wrt-9,* and *grl-23* and two receptor genes: *ptc-3* and *ptr-22* regulated by the DBL-1 pathway (Table 5). Interestingly, *wrt-4* and *ptc-3* were also listed in TGF-β dauer pathway regulated genes (Liu et al., 2004).

Major sperm protein

Major sperm proteins (MSPs) are sperm specific proteins that function as both extracellular signals to regulate ovulation rate and as a cytoskeleton to allow the mobility of sperm (Miller et al., 2004). Approximately 40 genes encode MSP in *C. elegans,* and their functions are redundant. 14 *msp* genes and three genes encoding conserved MSP domains were regulated by *dbl-1* (Table 5) (Table 8). However, none were regulated by *sma-9*. Interestingly, all these genes were upregulated. At the same time, *cav-1* (*C. elegans* caveolin) was downregulated (Table 5). This results are consistent with the idea that *cav-1* is a negative regulator of MSP (Miller et al., 2004). To date, a function of MSP in the DBL-1 pathway has not been reported. One possibility might be that those MSPs affect the brood size of the Sma mutants.

Germline genes

In *Drosophila,* genetic evidence has been provided that Bmp homologs *gbb* and *dpp* are essential for maintaining germline stem cells (GSC) in ovary and testis (Kawase et al., 2004; Song et al., 2004). The authors showed that Bmp signaling functions as a short-range signals for repressing *bam* (a gene promoting differentiation in GSC) in order to keep GSC at renewal and undifferentiated stage. Meanwhile, they reported that the Smad proteins Med and Mad physically interacts with the silencer elements of *bam in*

vitro, whose sequence shares high similarity with silencer elements of *brk.* However, in *C. elegans,* the functional relationship between TGF-β signaling and germline development is still a mystery. Our microarray data includes many genes affecting germline development, such as Lin-12/Notch pathway components (e.g. *hop-1, gld-1)*; P granule components (e.g. *glh-1, iff-1)*; sex determination genes (e.g. *fbf-2, gld-3)* and other genes (e.g. *cav-1, syp-3)* (Table 5). The majority of these meiosis genes were repressed.

It has been reported that Dpp signaling and Notch signaling have opposite effects in the fusion cell fate in dorsal branches of the *Drosophila* trachea (Steneberg et al., 1999). However, other reports suggest that TGF-β and Notch pathways function synergistically in the regulation of *hes1* and *hey1,* the immediate Notch signaling responsive target genes (Blokzijl et al., 2003; Dahlqvist et al., 2003) in cell lines. A physical interaction between intracellular domain of Notch1 and Smad3 or Smad1 was detected. In *C. elegans* the crosstalk between these two pathways is still poorly understood. We noticed that Lin-12/Notch pathway components existed in our array positives. Many gene partners in the pathway are redundant with each other, for example, *hop-1* and *sel-12* (Li and Greenwald, 1997); *gld-1* and *gld-2 (*Hansen et al., 2004*); fbf-1* and *fbf-2 (*Lamont et al., 2004*)* and so on. In our data, only one gene partner from these partnered genes changed their expression level. We do not completely understand the meaning of these phenomena.

The expression level of T27F2.4 was significantly upregulated in DBL-1 pathway mutants

In order to verify the genes identified in microarray analysis, eight candidates with high expression changes were tested. The candidate gene promoter regions (about 2kb) were fused with GFP reporter genes, and then injected into wild type and *sma-9* mutant background. Gene C47D12.1 and T27F2.4 showed consistence with the microarray results, but not the others. Perhaps, the promoter regions chosen for the other six genes failed to represent correct expression of the genes, or the predicted gene structure is incorrect.

Significantly, *T27F2.4::GFP* has a strong expression in *sma-9, dbl-1,* and *sma-6* (lf) mutants compared with wild type (Fig 9A-E). Both in wild type and DBL-1 pathway mutants, *T27F2.4::GFP* expression in neuronal cells was first observed at the L1 larval stage and then slowly disappeared after L4 (Fig 9B). In wild type, a very weak expression in intestine was observed only in adults. However, in the DBL-1 pathway mutants a weak expression in intestine occured during early larval stages and became much stronger after L4 stages (Fig 9A, C, D). Meanwhile, a detailed investigation in *sma-6(lf)* mutants showed a weaker signal compared with that in *sma-9* and *dbl-1* mutants (Fig 9E). Perhaps, in the absence of type I receptor *sma-6* the DBL-1 pathway still has basal activity through type II receptor *daf-4*.

To further investigate function of T27F2.4 in DBL-1 pathway, dsRNAi was performed in wild type and DBL-1 pathway (lf) mutants background (Table 9). By microinjecting T27F2.4 dsRNAi in *him-5* strain, the animals body size was slightly longer than control; their male tails contained a ray 3-4 fusion that was displayed in

ligand overexpression males (Fig 9F). Meanwhile, we used RNAi feeding vector to perform the RNAi in order to further quantify the phenotypes. However, all these animals were the same as the comparative control. This indicates that RNAi feeding method was not strong

Fig 9. Gene T27F2.4 might be a negative regulator in the DBL-1 pathway.

(A-E) T27F2.4 expression level was significantly changed in DBL-1 pathway (lf) mutants. In wild type (B), neuronal expression were observed at larval stages but disappeared in adulthood. The same expression pattern was displayed in DBL-1 pathway (lf) mutants (data not shown). T27F2.4 intestine expression was almost undetectable in wild type background. However, in *sma-9* background a high intestine expression was observed in adult (A); while a low intestine expression was detected in larval stages (C). *dbl-1(lf)* animal expressed a similar level of T27F2.4 in intestine (D) as in *sma-9*. In *sma-6 (lf)* mutants, the intestine expression is lower compared with *dbl-1* and *sma-9,* however it is still significantly higher than in wild type (E).

(F-H) T27F2.4 dsRNAi animals (G) are longer than control (H) and their males have a ray 3-4 fusions (F), indicating that it might be a negative regulator of DBL-1 pathway regulated body size and male tail morphogenesis.

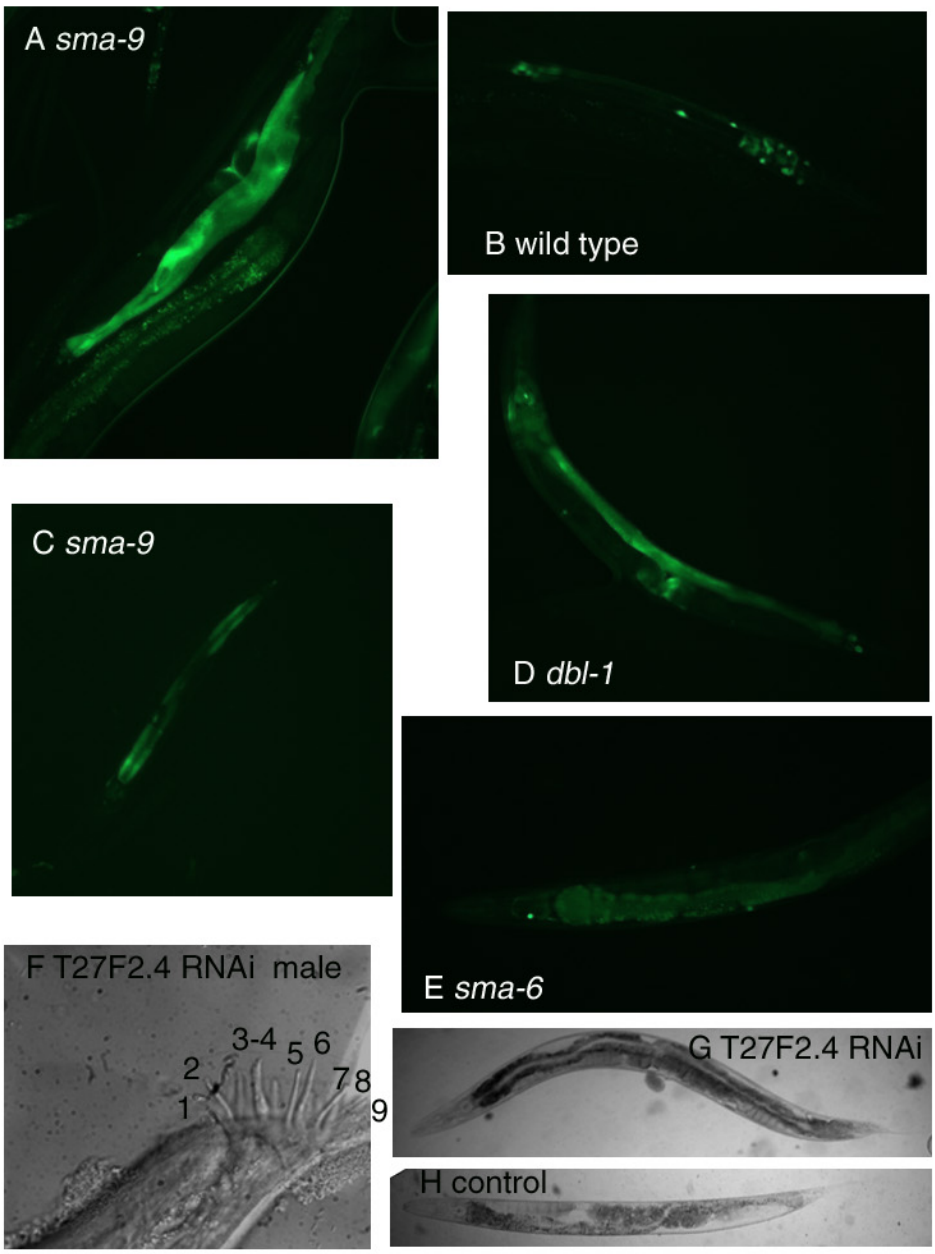

A *sma-9*
B wild type
C *sma-9*
D *dbl-1*
E *sma-6*
F T27F2.4 RNAi male
1
2
3-4
5
6
7
8
9
G T27F2.4 RNAi
H control

Table 9. T27F2.4 RNAi Mab phenotypes

Mutant	Fusion of Rays 3-4	Fusion of Rays 4-5	Fusion of Rays 6-7	Fusion of Rays 8-9	N (sides)
him-5(e1490)	0%	0%	0%	16%	74
sma-9 (*wk55)*	0%	0%	0%	53%	101
sma-3(wk30)	0%	27%	42%	15%	97
T27F2.4 RNAi in *him-5(e1490)*[1]	2.5%	0%	0%	8%	39
him-5(e1490)[2]	0%	0%	0%	23%	94
sma-9 (*wk55)*[2]	0%	0%	0%	64%	47
sma-3(wk30)[2]	0%	23%	51%	36%	69

1. RNAi performed by microinjection ;
2. RNAi performed by feeding vector.

enough to cause the phenotype. Although we did not observe any phenotype change in *sma-9* and *sma-3* background by feeding method, it did not rule out the possibility that T27F2.4 interacts with DBL-1 pathway components genetically.

T27F2.4 encodes a basic leucine zipper transcription factor homologous to human $p21^{SNFT}$ and mouse fosB. $p21^{SNFT}$ has been shown to repress gene transcription together with Jun, because of its capability of replacing Fos and dimerizing with Jun (Bower KE, et al 2002). However, FosB has been shown to stimulate cell proliferation by activating gene transcription upon restoring TGF-β signal (Deacu et al., 2004). In the case of *C. elegans,* T27F2.4 was repressed by the DBL-1 pathway and it may function as a negative regulator in the DBL-1 signaling regulated body size development and male tail morphogenesis.

DISCUSSION

sma-9 functions as a transcriptional repressor in the DBL-1 pathway

sma-9 encodes a zinc finger transcription factor, homologous to *Drosophila* Schnurri. Shn has been shown to repress target gene transcription in *brk* dependent and independent manners (Campbell and Tomlinson, 1999; Jazwinska et al., 1999; Marty et al., 2000; Pyrowolakis et al., 2004). However, some research supports that Shn activates some target genes (Torres-Vazquez et al., 2001). Since the *C. elegans* genome does not encode *brk,* several questions arise such as: how does *sma-9* function in the DBL-1 pathway? is it a transcriptional repressor or activator? Here, we show the first evidence that a transcriptional repressor can substitute for functional Shn protein *in vivo.*

Our genetic studies in using *VP-16-sma-9* and *en-sma-9* constructs showed that *en-sma-9* rescues both body size and mail tail defects in *sma-9* (lf) mutant (Fig 6, Table 2). Detailed examination of the rescued phenotype showed that En-SMA-9 gives rise to an overexpression phenotype of male ray 3-4 fusions both in *sma-9* mutants and wild type (Table 2), however it does not affect body size at all (data not shown). Male ray morphogenesis is likely more sensitive to the signal than body size development. Given that VP-16-SMA-9 did not cause a dominant negative effect, perhaps a minor function of *sma-9* is as a transcription activator. Taken together, our results support that *sma-9* mainly functions as a transcriptional repressor in DBL-1 pathway regulated body size development and male tail morphogenesis.

Meanwhile, we showed that *sma-9* N-terminal sequences have transcriptional repressor activity *in vitro* (Table 3). This helped to interpreted our negative results from yeast two-hybrid experiments. Furthermore, investigation of ARD II showed a transcriptional activator activity, which is also consistent with the yeast two-hybrid assay. Given that alternative splicing exists in *sma-9*, isoforms containing the N-terminus would act as transcriptional repressors, while some isoforms, for example *yk1237d01,* could function as a transcriptional activator in the pathway. There might be a balance between *sma-9* transcriptional repressor and activator activities. How precisely the two forces function *in vivo* is still a mystery. Thus, these *in vitro* results support our previous hypothesis.

The third line of evidence comes from our microarray analysis. Among genes regulated by *sma-9,* more genes are repressed than are activated, supporting that *sma-9* functions as a transcriptional repressor in the pathway in early larval stages (Fig 8).

Surprisingly, in *dbl-1* regulated gene 83% genes were repressed, suggesting that in early larval stage the DBL-1 pathway mostly represses target genes. One possibility is that the Smad complex represses the target genes directly, alternatively it recruits transcriptional repressors such as *sma-9,* or it turns on gene transcription of transcriptional repressor which then represses target genes. In our finding *sma-9* transcription was not regulated by the DBL-1 pathway itself, instead it might be recruited by Smad proteins in order to repress gene transcription.

sma-9 functions synergistically with _lin-49_ for male ray morphogenesis

The *lin-49* upstream region contains a SE sequence similar to SE from *Drosophila brk.* Meanwhile, the loss of function phenotype was similar with DBL-1 pathway overexpression phenotype. We ask whether *lin-49* might be a downstream target of *sma-9* or function as *C. elegans brk. lin-49* RNAi in *sma-9* (lf) background causes new fusions at rays 1-5, which was not seen in either *lin-49* RNAi or in *sma-9* loss of function mutants (Table 4). Thus it suggests that they function synergistically in ray 1-5 morphogenesis but not ray 6-9. Given *lin-49* RNAi in *sma-2* and *sma-3* mutants did not give rise to these new phenotypes and the male ray fusion pattern were the same as the mutants themselves, we propose that the synergistic function between *sma-9* and *lin-49* is DBL-1 pathway independent.

Genes regulated by _sma-9_ and _dbl-1_

It has been reported that the DBL-1 pathway regulates body size development through controlling collagens and chromosomal ploidy due to endoreplication in

hypodermal cells (Nystrom et al., 2002; Flemming et al., 2000; Morita et al., 2002). In our microarray profile, we see collagens genes, cell cycle control genes that might regulate endoreplication, and other genes affecting body size and growth e.g. Y54G9A.5 (Table 5, 6, 7, 8). Meanwhile, we see genes influencing innate immunity - another major function of the DBL-1 pathway, including a large number of C-type lectin genes, chemoreceptor genes; ABC transporters; and genes regulating dauer and aging etc (Table 5, 6, 7, 8). Also we identified genes of Hedgehog family. Genetic interactions between TGF-β pathways and Hedgehog signaling have been extensively studied in other organisms (Tabata and Takei, 2004), however little is known in *C. elegans*.

Interestingly, the gene T27F2.4 encoding a bZIP transcription factor showed significant expression level changes between wild type and Sma mutant in DNA microarray profiles (Table 8) and by a transcriptional GFP reporter (Fig 9). The major difference was intestine expression which is an important tissue for innate immunity. T27F2.4 knockout by dsRNAi demonstrated a *dbl-1* overexpression phenotype both in body size and male tail ray fusion, suggesting that it might act as a negative regulator in the DBL-1 pathway (Fig9; Table 5). Further analysis is necessary to address whether T27F2.4 is a target gene of the DBL-1 pathway or not and what is the role of T27F2.4 in the pathway.

Furthermore, other gene families in our microarray analysis include major sperm protein and germline genes (Table 5, 6, 7, 8). BMP pathway has been shown to influence germline development in *Drosophila* (Song et al., 2004). Analyzing function of these genes in DBL-1 pathway would enrich our knowledge.

Compared with the *dbl-1* regulated genes, the *sma-9* regulated genes mostly covered the similar gene categories but with certain differences (Table 5, 6, 7, 8). *sma-9* did not regulate MSP, ABC transporter, and cell cycle control genes. Only one collagen gene and a few chemoreceptors were regulated by *sma-9*. This restriction on gene regulation supports that *sma-9* functions specifically in the DBL-1 pathway: at early larval stages; specific male ray morphogenesis. We propose that in addition to *sma-9,* Smad proteins need other transcriptional cofactors to fulfill their complete functions.

Here we showed genetic and biochemical evidence that SMA-9 functions as a transcriptional repressor in DBL-1 pathway regulated body size development and male tail morphogenesis, which is supported by gene profiles. Furthermore, different gene categories were identified from microarray analysis including genes regulating characterized function e.g. body size and innate immunity, as well as genes with new functions in DBL-1 pathway e.g. hedgehog signaling and germline development. Studies on these new functions would enrich our knowledge of DBL-1/BMP/TGF-β pathway.

CONTRIBUTIONS: This work was accomplished with technical assistance from technician Ling Yu. She helped to clone the T27F2.4 promoter into an RNAi vector and microinject the dsRNA into wide type animals; she also microinjected T27F2.4 – GFP PCR product and C47D12.1 construct into wild type background. In the *lin-49* RNAi experiments, she helped transforming *lin-49* construct into feeding bacterial strains and placed the Sma mutant animals on the plates. I thank Dr. Gary Struhl for the VP-16 and En vectors; Dr. Sergei Kuchin for the yeast vectors and strains; Dr. Iva Greenwald for

helpful discussions; Dr. Agnes Viale (Genomics Core Facility, MSKCC) for microarray analysis; Dr. Xinzhu Deng and Dr. Richard Kolesnick for allowing me to use γ-irradiation for integration.

CONCLUSION

TGF-β signaling pathways are critical for animal development. In our lab, we use *C. elegans* as a model system to study the pathway, in particular the DBL-1/BMP/ TGF-β pathway. Loss of function of the DBL-1 pathway causes a small body size (Sma) and abnormal male tail (Mab). From a forward genetic screen, mutants in a new gene *sma-9* were isolated.

sma-9 mutants display a similar body length and male tail defect as other DBL-1 pathway components do, which attracted me to clone this gene. It turned out that SMA-9 is a transcription factor, homologous to *Drosophila* Schnurri that is required for Dpp/ TGF-β pathway. The expression pattern of *sma-9* showed that it widely exists in most worm tissues: intestine, pharynx, vulva, ventral nerve cord, seam cells, spermatheca, etc. Its expression overlaps with other DBL-1 pathway components e.g. *dbl-1, sma-2, sma-3, sma-4* etc. Thus, these results support that *sma-9* functions in the DBL-1 pathway. Meanwhile, genetic studies put *sma-9* downstream of *dbl-1*, and suggest that it functions at early larval stages in body size development.

Hyp7 has been shown to be a critical tissue in DBL-1 pathway regulated body size development. However based on reporter constructs, *sma-9* is not transcribed here. Question arises as to: how does *sma-9* regulate body size via the DBL-1 pathway? Given that during larval development seam cells fuse into hyp7, it is possible that the fusion would bring *sma-9* message into hyp-7 and then proceed with *sma-9* translation. Consistent with this, Dr. Jun Liu's lab used SMA-9 antibody to perform immunohistochemistry, and showed that SMA-9 is located at hyp7. Thus, their results proved the above hypothesis. They also showed that the protein is localized in the nuclei,

supporting that SMA-9 acts as a transcription factor. All these results suggest that the seam cell is another important tissue of DBL-1 pathway regulated body size development.

Meanwhile, we observed a complex alternative splicing in *sma-9*. The transcriptional expression patterns using different lengths of promoter regions demonstrated that different *sma-9* isoforms may be transcribed in different tissues. It is possible that different isoforms perform various functions in the DBL-1 pathway or other completely different functions, which needs further investigation.

The *sma-9* homolog Schnurri (Shn) play a critical role in *Drosophila* Dpp pathway. Several researches suggest that Shn functions as a transcriptional repressor. However, other reports indicate that Shn could be a transcriptional activator. In order to address whether *sma-9* is a transcriptional repressor or activator, new artificial *sma-9* constructs were generated that would act as a repressor SMA-9 and activator SMA-9 respectively.

Genetic studies of these two *sma-9* constructs showed that repressor SMA-9, but not activator SMA-9, partially rescued the *sma-9* mutants phenotypes and gave rise to *dbl-1* overexpresson male tail phenotypes. *In vitro* transcriptional activity assay demonstrated that SMA-9 N-terminal regions have transcriptional repressor activity. Thus, these results indicate *sma-9* acts as a transcriptional repressor in the DBL-1 pathway.

In order to find target genes of the DBL-1 pathway, microarray techniques were used. *dbl-1* and *sma-9* mutants at L2 stage were chosen to perform the assay. From the DNA microarray analysis, there were more genes being repressed than being activated by *dbl-1* and *sma-9*. This is not surprising for *sma-9* since it is a transcriptional repressor,

but it is surprising for *dbl-1,* and indicates that DBL-1 signaling represses target genes in general. Further studies are absolutely necessary in order to fully understand the DBL-1/ TGF-β signaling.

Genes pulled out from the DNA microarray include a large variety of genes, for example, collagens, cell cycle genes, dauer and aging genes, and so on. Interestingly, many genes related to innate immunity were identified, which is one of the major functions of the DBL-1 pathway. Hedgehog genes were also in the lists suggesting a genetic interaction between these two signaling pathways. In addition, many germline genes were found, which indicates that the DBL-1 pathway may function in germline development. Investigating genes from microarray profile would give us clues of new functions of the DBL-1 pathway.

The DBL-1 pathway has been demonstrated to regulate gene T27F2.4 transcription in this study. Mammalian homologs of T27F2.4 have multiple functions including function in the TGF-β pathway, which attracts me to further investigate this gene. *C. elegans* T27F2.4 knockdown phenotypes displayed similar Lon and Mab phenotypes as other DBL-1 pathway component do, suggesting that it might be a negative regulator of this pathway. However loss of function alleles are not available, which make it hard to precisely investigate genetic interactions of T27F2.4 and the DBL-1 pathway. Since a strong expression level change was observed between DBL-1 pathway mutants and wild type animals, it would be an excellent marker for the DBL-1 pathway.

The gene *sma-9* has various functions during worm development. In addition to the DBL-1 pathway, this research proposed that it may regulate male ray morphogeneis

of ray 1-5. The double knockdown of *sma-9* and *lin-49* causes severe male ray 1-5 fusions, which is unlike the single mutants and other DBL-1 pathway components. Since *lin-49* encodes a transcription factor, the protein LIN-49 might interact with SMA-9. It will be interesting to know how they interact with each other.

In the future, identifying more target genes of the DBL-1 pathway will be one of the major jobs that researchers should do. In addition, investigating new functions of the pathway is absolutely necessary, in particular germ line development, hedgelog signaling, innate immunity etc. The mystery needs to be solved as to how precisely does *sma-9* function in the pathway. For example, the SMA-9 binding element is still unknown; how and where Smads interact with SMA-9; what are each SMA-9 isoform's function; and so on. It will be of great interest to know all these, plus new *sma-9* functions during animal development.

Table 6. Genes specifically regulated by *dbl-1* in DNA microarray analysis

Genes upregulated in *dbl-1* background	
B0205.2	srz-85: Serpentine Receptor, class Z; 7-transmembrane receptor
B0218.6	C-type lectin
B0272.4	Enoyl-CoA hydratase/isomerase
B0304.6	sra-33: Serpentine Receptor, class A (alpha)
B0334.9	Ribosome-inactivating protein
B0511.1	fkb-7: FK506-Binding protein family; FKBP-type peptidyl-prolyl cis-trans isomerase; homologous to mammalian FK506-binding proteins
B0511.5	Cuticulin precursor; Zona pellucida-like domain & Endoglin/CD105 antigen
C01F6.2	unknown function
C01G10.5	unknown function
C01G5.4	WSN domain
C01G6.7	Acyl-CoA synthetase ; AMP-binding enzyme domain
C01G6.9	unknown function
C02F5.11	tsp-2: Tetraspanin family
C04C3.4	G2/Mitotic-specific cyclin A
C04E12.1	DUF40
C04F6.1	vit-5: a precursor of the lipid-binding protein related to ApoB-100
C04G2.1	Uncharacterized protein with conserved cysteine; Transthyretin-like family domain
C04G2.4	msp-36: major sperm protein
C05D9.9	Predicted membrane proteins, contain hemolysin III domain, Immunoglobulin-like domain
C05E7.4	Rhodopsin-like GPCR superfamily domain
C06A8.6	Protein phosphatase 1, regulatory subunit, and related proteins
C06G3.12	Transposon-encoded proteins with TYA, reverse transcriptase, integrase domains in various combinations
C08A9.1	sod-3: SOD (superoxide dismutase); a iron/manganese superoxide dismutase
C08F11.11	DUF1505
C08F8.5	Predicted transposase; containing an F-box
C09D4.3	Casein kinase (serine/threonine/tyrosine protein kinase)
C09G1.2	unknown function
C10G8.4	Trypsin Inhibitor like cysteine rich domain
C12D8.9	unknown function
C13D9.1	srr-6: Serpentine Receptor, class R
C15C8.3	Aspartyl protease
C15H9.10	containing Phosphopantetheine attachment site
C15H9.8	prx-3: Peroxisome assembly factor; orthologous to the human gene peroxisomal biogenesis factor 3
C16C8.4	Ubiquitin and ubiquitin-like proteins
C16C8.5	Ubiquitin and ubiquitin-like proteins
C17B7.4	Predicted secreted cysteine rich protein found only in C.elegans; DUF19
C17C3.1	Acyl-CoA thioesterase
C17C3.10	hlh-27: Helix Loop Helix; Helix-loop-helix DNA-binding domain
C17G1.2	unknown function

C17H12.8	DUF141
C18A11.1	unknown function
C18E3.5	U5 snRNP-specific protein-like factor and related proteins
C24A11.1	unknown function
C24A3.3	unknown function
C24A3.4	Predicted L-carnitine dehydratase/alpha-methylacyl-CoA racemase
C24D10.8	unknown function
C24H10.3	unknown function
C26C6.6	Adaptor protein Enigma and related PDZ-LIM proteins
C27D6.10	srb-1: Serpentine Receptor, class B (beta)
C27D6.3	unknown function
C31B8.13	srh-247: Serpentine Receptor, class H; Predicted olfactory G-protein coupled receptor; 7TM chemoreceptor
C31H1.5	domain of Phospholipase A2
C32H11.10	dod-21: Downstream Of DAF-16 (regulated by DAF-16); DUF141
C33A12.13	sru-2: Serpentine Receptor, class U; DUF215
C33D3.3	unknown function
C33F10.11	unknown function
C34F11.4	msp-50: Major sperm protein
C35A5.2	UDP-glucuronosyl and UDP-glucosyl transferase
C35A5.6	unknown function
C35B8.1	col-175: Cuticle collagen; Collagens type IV and type XIII
C35D10.11	MSP domain
C36C5.12	Predicted secreted cysteine rich protein found only in C.elegans; DUF19
C36C5.14	Predicted secreted cysteine rich protein found only in C.elegans; DUF19
C37C3.10	unknown function
C40H5.1	unknown function
C43H8.3	unknown function
C44B12.8	srx-14: Serpentine Receptor, class X; 7-transmembrane receptor ; Rhodopsin-like GPCR superfamily
C44B7.5	C44B7.5; Predicted membrane protein, contains DoH and Cytochrome b-561/ferric reductase transmembrane domains
C44B7.9	pmp-2: Peroxisomal Membrane Protein related; Long-chain acyl-CoA transporter, ABC superfamily (involved in peroxisome organization and biogenesis)
C44C10.3	C44C10.3: Synaptic vesicle transporter SVOP and related transporters
C45B11.3	dhs-18: Dehydrogenases, Short chain
C45B2.1	unknown function
C45B2.2	unknown function
C48B4.6	unknown function
C48E7.8	Predicted acyltransferase
C50D2.8	Spliceosome subunit
C50H11.9	str-244: 7-transmembrane olfactory receptor
C52A10.1	Carboxylesterase and related proteins
C52B11.5	GTPase Rab5/YPT51 and related small G protein superfamily GTPases; Ras family domain
C52E2.6	Predicted transposase
C53C11.4	unknown function

CEESL04 clone	multiple target
D1054.10	unknown function
D1086.6	unknown function
D2021.4	Acyl-CoA thioesterase
D2096.1	Mitochondrial import inner membrane translocase, subunit TIM17
E02A10.2	grl-23: GRound-Like (grd related); hedgehog-related
E04F6.3	Peroxisomal multifunctional beta-oxidation protein and related enzymes; MaoC like domain
E04F6.8	unknown function
F01D4.2	UDP-glucuronosyl and UDP-glucosyl transferase
F01F1.14	unknown function
F01G10.2	ech-8: Enoyl-CoA Hydratase
F02H6.6	Zinc metalloprotein; type I Thrombospondin domain
F08B12.2	prx-12: Peroxisome assembly factor: Predicted E3 ubiquitin ligase involved in peroxisome organization
F08E10.7	Defense-related protein containing SCP domain; Allergen V5/Tpx-1 related domain
F08F1.9	G2/Mitotic-specific cyclin A
F08F8.5	unknown function
F09C12.7	msp-74: Major sperm protein
F10D2.11	UDP-glucuronosyl and UDP-glucosyl transferase
F10D2.4	str-112: 7-transmembrane olfactory receptor; member of the ODR-10 family of G-protein coupled receptors.
F10D2.6	UDP-glucuronosyl and UDP-glucosyl transferase
F11A6.2	homolog of human Phospholipid scramblase
F11G11.8	F11G11.8, DUF780
F12A10.1	unknown function
F13A7.7	unknown function
F14B8.4	unknown function
F14D7.7	unknown function
F14F8.6	srw-44: 7-transmembrane olfactory receptor
F14H3.3	unknown function
F16B4.4	unknown function
F17A9.5	NADH:flavin oxidoreductase/12-oxophytodienoate reductase
F18A1.7	unknown function
F20G2.5	DUF141
F21E9.3	Uncharacterized protein with conserved cysteine; Transthyretin-like family
F22A3.4	unknown function
F22B5.5	DUF1248 domain
F22D6.8	unknown function
F22F4.4	unknown function
F23H11.7	unknown function
F25D1.5	short chain dehydrogenase or Reductases
F25E2.2	ShTK domain
F25H5.8	Stress responsive protein
F26A3.3	ego-1: Enhancer of Glp-One (glp-1); RNA-directed RNA polymerase essential for germ-line development

F26G1.7	msp-3: Major sperm protein
F29A7.7	C-type lectin
F32B6.6	msp-77: Major sperm protein
F33D11.1	Adaptor protein Enigma and related PDZ-LIM proteins
F34D6.1	containing an F-box, a motif mediate protein-protein interactions
F35E12.5	F35E12.5; DUF141
F35F10.13	F35F10.13:DUF19
F35F10.7	Predicted pyroglutamyl peptidase ; DUF976
F36H1.5	unknown function
F36H12.15	unknown function
F39H2.4	syp-3: Synapsis in meiosis abnormal
F40F4.3	lbp-1: Lipid Binding Protein; Fatty acid-binding protein FABP
F42G10.1	Encodes a neprilysin, a thermolysin-like zinc metallopeptidases, F42G10.1 has no clear orthologs in other organisms.
F43D9.4	sip-1: Stress Induced Protein; Alpha crystallins
F44A6.4	unknown function
F44D12.7	MSP domain (major sperm protein domain)
F45C12.16	Predicted receptor: EGFR- (epidermal growth factor receptor) activity
F45E4.8	nlp-20: Neuropeptide-Like Protein; predicted neuropeptide of the FAFA family;
F45F2.11	unknown function
F47B8.3	Glutaredoxin-related protein
F49C5.7	unknown function
F49E8.2	Zn-finger, C2H2 type
F52D2.5	unknown function
F52E1.1	pos-1:Posterior Segregation gene; CCCH-type Zn-finger protein;
F52E1.5	unknown function
F52H3.4	unknown function
F53A9.1	containing domain of Peptidase M, neutral zinc metallopeptidases
F53A9.2	domains of Peptidase M, neutral zinc metallopeptidases
F53A9.7	unknown function
F53B6.8	unknown function
F53C11.3	short chain dehydrogenase & Reductases domain
F53E10.4	unknown function
F53F1.7	srv-8: Serpentine Receptor, class V
F54F3.4	short chain dehydrogenase or Reductases
F54H12.4	unknown function
F55A11.8	unknown function
F56C3.9	unknown function
F56D1.7	daz-1: human DAZ (Deleted in Azoospermia) homolog; containing an RNA recognition motif that is required for the progression of meiosis during oogenesis
F56H6.3	unknown function
F56H9.3	gpa-8: G-protein alpha subunit (small G protein superfamily) of heterotrimeric GTPases;
F57G4.5	unknown function
F58A6.1	Enoyl-CoA isomerase
F58E6.7	unknown function
F58F6.6	Membrane protein; DUF621

F58G4.4	Proteins containing Ca2+-binding EGF-like domains
F58H1.6	EGF-like domain
F59C6.6	nlp-4: Neuropeptide-Like Protein
F59D8.A	vit-3: a precursor of the lipid-binding protein related to ApoB-100
F59E11.7	unknown function
F59F4.1	Acyl-CoA oxidase
F59H6.11	contains BTB/POZ domain & MATH domain
F59H6.9	contains BTB/POZ domain
H04M03.6	srv-19: Serpentine Receptor, class V; DUF280
H04M03.8	srv-21: Serpentine Receptor, class V; DUF280
H09G03.1	unknown function
H11L12.1	unknown function
H12D21.1	unknown function
H12I19.5A	H12I19.5A
H23N18.1	UDP-glucuronosyl and UDP-glucosyl transferase
H23N18.3	UDP-glucuronosyl and UDP-glucosyl transferase
K01D12.11	cdr-4: Cadmium Responsive;
K01H12.2	Mitochondrial ADP/ATP carrier proteins
K02E10.6	CX module
K03B4.5	srx-77: Serpentine Receptor, class X; G protein coupled protein
K04F1.9	unknown function
K05B2.4	Peroxisomal long chain acyl-CoA thioesterase I/predicted bile acid-CoA-amino acid N-acyltransferase
K06C4.9	7 transmembrane receptor (rhodopsin family)
K06H6.1	DUF268
K07A1.13	unknown function
K07F5.1	msp-81: Major sperm protein
K07F5.2	msp-10: Major sperm protein
K07F5.5	DUF 780
K07H8.6	vit-6: Vitellogenin structural genes (yolk protein genes)
K08B4.6	cli-2: Cystatin-LIke protease inhibitor
K08D8.5	DUF141 domain
K08F4.8	msp-38: Major sperm protein
K09H11.7	p-Nitrophenyl phosphatase;
K09H9.4	reverse transcriptase
K10B2.3	C-type lectin
K10G4.3	unknown function
K11G12.7	acr-3: AcetylCholine Receptor; encodes a non-alpha subunit of the nicotinic acetylcholine receptor (nAChR) superfamily; member of the UNC-29-like group of nAChR subunits;
M01D1.5	unknown function
M02H5.4	Hormone receptors; Zn-finger, C4-type steroid receptor
M60.7	Ankyrin repeat
M70.1	WSN domain
R02E12.6	unknown function
R03A10.5	Phosphatidylinositol transfer protein SEC14 and related proteins
R03H4.4	srt-18: Serpentine Receptor, class T; 7-transmembrane receptor

R05F9.13	msp-31: Major sperm protein
R05F9.5B	gst-9: Glutathione S-Transferase
R05F9.8	msp-33: Major sperm protein
R05H10.1	unknown function
R07B1.2	lec-7: Galectin, galactose-binding lectin
R07B1.2	lec-7: Galectin, galactose-binding lectin
R07B7.13	Nuclear hormone receptor
R09B5.9	cnc-4: Caenacin (Caenorhabditis bacteriocin)
R09F10.6	srh-11: Serpentine Receptor, class H; Predicted olfactory G-protein coupled receptor; 7TM chemoreceptor
R10D12.9	Multitransmembrane protein
R11A5.5	unknown function
R11D1.3	Small secreted protein with conserved cysteines
R11G11.3	Proteinase inhibitor I4, serpin
R13H4.3	Lysosomal & prostatic acid phosphatases
T01B11.4	Mitochondrial ADP/ATP carrier proteins
T01C8.2	unknown function
T02H6.3	unknown function
T03F7.4	srh-47: Serpentine Receptor, class H; Predicted olfactory G-protein coupled receptor
T03G11.2	sru-48: Serpentine Receptor, class U; DUF215
T03G11.8	zig-6: two (Zwei) IG-domain protein; membrane-associated protein containing two C2-type immunoglobulin domains and a GPI (glycosylated phosphatidylinositol) anchor site, but no transmembrane domain;
T04G9.7	unknown function
T04H1.9	tbb-6: Beta tubulin
T05A10.3	Uncharacterized protein with conserved cysteine; Transthyretin-like family domain
T05A10.6	Allergen V5/Tpx-1 related domain
T05A7.9	unknown function
T06C12.5	containing an F-box and an FTH/DUF38 motif, mediating protein-protein interactions;
T07F10.5	G2/Mitotic-specific cyclin A
T07H8.1	srx-16: Serpentine Receptor, class X; 7-transmembrane receptor
T08A9.8	spp-4: Saposin-like Protein family
T09B4.6	unknown function
T09D3.2	srg-28: Serpentine Receptor, class G (gamma)
T10B5.1	unknown function
T11F1.1	srw-60: Serpentine Receptor, class W; 7-transmembrane olfactory receptor
T13C5.1	cytochrome P450 of the CYP2 subfamily;
T13F2.10	msp-79: major sperm protein
T13F2.11	msp-78: major sperm protein
T14G8.2	Cyclin-like F-box
T15B12.2	Casein kinase (serine/threonine/tyrosine protein kinase)
T15B7.4	col-142: Collagens (type IV and type XIII)
T15B7.5	col-141: Collagens (type IV and type XIII)
T15B7.6	unknown function
T19C4.5	Globin-like domain

T19H12.10	UDP-glucuronosyl and UDP-glucosyl transferase
T20D4.4	Predicted peptide:N-glycanase
T20D4.7	Thioredoxin, nucleoredoxin and related proteins
T20H12.2	unknown function
T21D12.6	unknown function
T22F3.4	rpl-11.1: 60S ribosomal protein L11
T22G5.7	spp-12: Saposin-like Protein family
T23B7.1	DUF780
T23F11.2	unknown function
T23F11.5	srg-13: Serpentine Receptor, class G (gamma)
T23F2.3	stress responsive protein
T24A6.2	unknown function
T24B8.5	Secreted surface protein
T25C12.2	spp-9: Saposin-like Protein family
T25D1.3	F-box domain, mediating protein-protein interactions either with homologs of yeast Skp-1p or with other proteins
T26E4.6	unknown function
T27C5.10	7-transmembrane olfactory receptor
T28A11.20	neprilysin-thermolysin-like zinc metallopeptidases
T28A11.3	Predicted secreted cysteine rich protein found only in C.elegans; DUF19
T28A11.5	Predicted secreted cysteine rich protein found only in C.elegans; DUF19
T28A11.6	unknown function
T28A8.5	DUF545 domain
T28H10.2	Uncharacterized conserved protein; ShTK domain
W02B3.6	unknown function
W04B5.1	unknown function
W05G11.3	col-88: Collagens (type IV and type XIII)
W06B3.1	Nicotinamide mononucleotide adenylyl transferase
W07B8.1	Cysteine proteinase Cathepsin L
W09B7.2	unknown function
Y105E8B.B	ech-7: Enoyl-CoA Hydratase
Y106G6D.4	Glycogen synthase kinase-3
Y106G6D.6	unknown function
Y116A8C.41	7transmembrane olfactory receptor
Y11D7A.6	Reverse transcriptase
Y15E3A.3	unknown function
Y22F5A.6	lys-3: N-acetylmuraminidase/lysozyme
Y38C9B.3	Predicted receptor
Y38E10A.7	Triacylglycerol lipase
Y39C12A.8	dnj-26: DnaJ domain (prokaryotic heat shock protein); Molecular chaperone (DnaJ superfamily)
Y39G10BM.E	Actin filament-binding protein Afadin
Y43C5A.1	DUF780
Y43C5A.3	unknown function
Y45F10C.4	DUF1505
Y45F10D.6	unknown function

Y46C8AL.B	Y46C8AL.2
Y47D3A.13	unknown function
Y47G6A.14	Zinc finger, C3HC4 type (RING finger)
Y48A6B.9	Zn2+-binding dehydrogenase (nuclear receptor binding factor-1)
Y49G5A.1	Serine proteinase inhibitor (KU family)
Y4C6B.2	Amino acid transporters
Y51A2C.1	unknown function
Y52E8A.4	Uncharacterized conserved protein; DUF895 domain
Y53F4B.26	unknown function
Y54F10AM.G	Y54F10AM.7
Y54G9A.5	Y54G9A.5; negative regulation of body size since RNAi animal are long
Y57A10C.6	Peroxisomal 3-ketoacyl-CoA-thiolase P-44/SCP2
Y57G7A.7	srz-64: Serpentine Receptor, class Z
Y59A8B.4	srh-134: Serpentine Receptor, class H; Predicted olfactory G-protein coupled receptor; 7TM chemoreceptor
Y73C8B.1	DUF1057 & Esterase/lipase/thioesterase domain
Y73C8B.3	DUF1057 & Esterase/lipase/thioesterase domain
Y75B8A.4	Mitochondrial ATP-dependent protease PIM1/LON
Y76G2A.2	unknown function
Y80D3A.7	ptr-22: Patched Related family
Y82E9BL.11	F-box domain
Y87G2A.2	Acyl-CoA thioesterase
yk116f3.3	T23D8.7: Translation initiation factor 2C (eIf-2C) and related proteins
yk119g10.3	Y77E11A.12: Collagens (type IV and type XIII)
yk122a10.5	multiple target
yk122f5.5	multiple target
yk148h7.3	Y62H9A.6
yk14h8.3	fbf-2: RNA-binding protein ; negatively regulate the activity of gld-1 mRNA
yk163a8.3	Y73C8B.3: DUF1057
yk198f7.3	cej-1: a direct target of GLD-1; cell junction protein;
yk207e6.3	grl-25: Ground-Like (grd related)
yk244c3.3	bli-1: Blistered cuticle Collagens
yk28d11.3	K02B9.1
yk295d7.3	K07B1.8
yk308a1.3	K07A1.6
yk324f5.3	T05A7.6: Casein kinase (serine/threonine/tyrosine protein kinase)
yk357e12.3	C29F7.6: DNA-binding protein jumonji/RBP2/SMCY, contains JmjC domain
yk437a5.3	multiple target
yk496f10.3	C41G7.3: Peripheral-type benzodiazepine receptor and related proteins
yk502a2.5	R09B3.2: mRNA cleavage and polyadenylation factor I complex, subunit RNA15
yk504b11.5	F58F9.7: Pristanoyl-CoA/acyl-CoA oxidase
yk545g6.3	H02I12.1: Chitin binding Peritrophin-A domain
yk557f1.3	multiple target
yk559h7.5	F23F12.12
yk564f8.3	H28G03.2
yk58a2.3	col-120: Collagens (type IV and type XIII)

yk609d3.3	gld-3: defective in Germ Line Development
yk645h12.3	F18E3.11
yk679d2.5	C17C3.12: Short-chain acyl-CoA dehydrogenase
yk705f9.3	K02C4.2
yk706e7.3	F07H5.10: Nuclear 5'-3' exoribonuclease-interacting protein, Rai1p
yk710d11.3	K02G10.3: Uncharacterized conserved protein
yk711h1.3	F59E11.7
yk713f8.3	rrn-3.1: Ribosomal RNA
yk715a12.3	na
yk724b1.5	B0410.3:
yk736d11.5	K07C5.5: Soluble epoxide hydrolase
yk737d7.3	F10E9.12: Copper amine oxidase & Plant inhibitor of proteinase and amylase domain
ZC178.1	unknown function
ZC373.5	Class II Aldolase and Adducin N-terminal domain
ZC373.7	col-176: Collagens (type IV and type XIII)
ZC410.1	nhr-11:Nuclear Hormone Receptor family
ZC410.6	unknown function
ZC443.5	UDP-glucuronosyl and UDP-glucosyl transferase
ZC455.11	DUF40
ZC64.3	ceh-18: C.Elegans Homeobox gene; ceh-18 encodes a POU-class homeodomain transcription factor;
ZK105.1	Predicted secreted cysteine rich protein found only in C.elegans; DUF-19 domain
ZK1248.17	MSP domain
ZK1248.4	Major sperm protein
ZK1251.1	Histone 2A
ZK1251.6	msp-76: Major sperm protein
ZK131.4	his-10: Histone H4
ZK402.3	DUF545
ZK512.7	unknown function
ZK546.6	msp-152: Major sperm protein
ZK550.6	Peroxisomal phytanoyl-CoA hydroxylase; ortholog of the human gene Phytanoyl-CoA hydroxylase which when mutated leads to Refsum disease
ZK697.8	Transthyretin and related proteins
ZK75.1	ins-4: Insulin-like peptides
ZK829.1	short chain dehydrogenase or Reductases
ZK863.2	col-37: Collagens (type IV and type XIII)
ZK993.3	unknown function

Genes downregulated in *dbl-1* background

B0344.2	wrt-9: Warthog (hedgehog-like family); Sonic hedgehog
C03A7.4	pqn-5: Prion-like-(Q/N-rich)-domain-bearing protein
C03H5.1	C-type lectin
C05C10.4	Lysosomal & prostatic acid phosphatases
C06G3.6	Uncharacterized conserved protein, contains ZZ-type Zn-finger
C09E8.3	DUF644

C14B9.4	plk-1: Polo Kinase; serine/threonine polo-like kinase; PLK-1 is required for meiotic nuclear envelope breakdown, polar body formation and extrusion, and proper chromosome segregation during meiosis, and germline development;
C29F3.5	C-type lectin
C32H11.11	Predicted alpha-helical protein ; DUF148
F07E5.7	Extracellular protein with cysteine rich structures ; DUF130
F07H5.9	Lysosomal & prostatic acid phosphatases
F08H9.7	C-type lectin
F11C7.3	vap-1: Venom-Allergen-like Protein; Defense-related protein containing SCP domain extracellular protein
F13H8.5	unknown function
F15E11.13	unknown function
F16H6.1	C-type lectin
F20B10.3	unknown function
F29C12.1	pqn-32: Prion-like-(Q/N-rich)-domain-bearing protein
F30H5.3	Serine proteinase inhibitor (KU family)
F32D1.10	mcm-7: DNA replication licensing factor, MCM7 component
F33D4.6A	F33D4.6A
F37B4.2	ifc-1: Intermediate Filament C; Nuclear envelope protein lamin; nonessential intermediate filament protein; predicted to function as a structural component of the cytoskeleton;
F57F4.3	gfi-1: Gei-4 Interacting protein; contains 21 ET modules; interacts with unc-68 in yeast two-hybrid assays.
K01A2.3	unknown function
K01A2.4	unknown function
K01D12.5	unknown function
K04H4.2A	Fibrillins and related proteins containing Ca2+-binding EGF-like domains
K07E8.3	Single-stranded DNA-binding replication protein A (RPA), large (70 kD) subunit and related ssDNA-binding proteins
K08D10.10	7-transmembrane receptor
K08E7.9	pgp-1: P-Glycoprotein related; transmembrane protein that is a member of the P-glycoprotein subclass of the ATP-binding cassette (ABC) transporter superfamily; along with PGP-3, PGP-1 is required for defense against the pathogenic Pseudomonas aeruginosa strain PA14
K11G9.4	egl-46: Egg Laying defective; transcription factor; a TFIIA-like zinc finger protein family
R02F11.1	Fibrillins and related proteins containing Ca2+-binding EGF-like domains
R09H10.5	EGF-like subtype 2 domain; Type I EGF domain
R11G11.6	Predicted alpha-helical protein ; DUF148
T01B7.7	rol-6: cuticle collagen (type IV and type XIII); ROL-6 interacts with SQT-1, a closely related cuticle collagen, expressed all stages from L1 to adult molts
T01C4.1	Chitinase
T01D3.6B	von Willebrand factor and related coagulation proteins ; Ficolin and related extracellular proteins
T05B4.3	Secreted surface protein ; ShTK domain
T05E12.6	DUF141
T05G5.3	cdk-1: Cyclin-Dependent Kinase family: orthologous to and functionally interchangeable with CDC28 from S. cerevisiae; required for cell-cycle progression through M phase in both meiosis and mitosis
T10B10.1	col-41: Collagens (type IV and type XIII)

T10H9.5	pmp-5: Peroxisomal Membrane Protein related
T19D12.1	unknown function
T21E8.1	pgp-6: P-GlycoProtein related; Multidrug/pheromone exporter, ABC superfamily
T23F1.6	pqn-71: Prion-like-(Q/N-rich)-domain-bearing protein
T23F2.1	Glycosyltransferase
T23G5.1	rnr-1: Ribonucleotide Reductase alpha subunit
VZK822L.1	fat-6: delta-9 fatty acid desaturase
W08A12.3	W08A12.3
Y105C5A.13	unknown function
Y110A2AL.8	ptc-3: Patched family
Y38H6C.16	C-8,7 sterol isomerase
Y48B6A.8	ace-3: abnormal Acetylcholinesterase;
Y57G11C.24C	eps-8: Eps (human endocytosis) related gene;
yk130f7.5	dhs-6: Dehydrogenases, Short chain
yk226c12.3	B0507.1: EGF-like domain
yk291a6.3	pqn-13: Prion-like-(Q/N-rich)-domain-bearing protein
yk394b9.3	lad-2: L1 Cam adhesion molecule homolog; Neural cell adhesion molecule L1
yk433c7.3	T06A4.1: Zinc carboxypeptidase
yk554b3.5	dod-3: regulated by DAF-16
yk63g3.3	unknown function
yk6e5.3	unknown function
yk716b10.5	T07D10.4: C-type lectin
yk728g5.3	C40H1.7: Predicted lipase
ZC123.1	unknown function
ZK1025.5	unknown function
ZK1025.7	DUF23
ZK1151.3	unknown function
ZK180.6	unknown function
ZK20.6	nep-1: Neprilysin metallopeptidase family; M13 family peptidase ; thermolysin-like zinc metallopeptidases,
ZK973.6	anc-1: abnormal nuclear Anchorage; nuclear envelope localization domain (the KASH domain) and an actin-binding domain; ANC-1 affects the positioning of nuclei and mitochondria within the cytoskeleton;
ZK973.7	unknown function

Table 7 Genes specifically regulated by *sma-9*

Genes upregulated in *sma-9* background	
B0213.2	nlp-27: Neuropeptide-Like Protein
B0213.3	nlp-28: Neuropeptide-Like Protein
B0213.4	nlp-29: Neuropeptide-Like Protein
B0348.2	domain of LPS-induced TNF (tumor necrosis factor) alpha factor
C01G8.6	unknown funtion
C03C11.2	fog-3: Feminization Of Germline; Anti-proliferation factor BTG1/TOB
C04E6.7	Predicted hydrolase (HAD superfamily)
C05E4.1	srp-2: encodes an ovalbumin-like serpin (ov-serpin) member of the serine protease inhibitor superfamily;
C06H5.2	containing an F-box and an FTH/DUF38 motif, both mediate protein-protein interaction.
C08E3.13	unknown funtion
C09G5.5	col-80: Collagens (type IV and type XIII)
C17H1.7	unknown funtion
C18H7.1	von Willebrand factor and related coagulation proteins
C23G10.11	unknown funtion
C27H5.4	unknown funtion
C29F9.12	unknown funtion
C31B8.4	unknown funtion
C32H11.3	DUF141
C32H11.4	DUF141
C34C12.7	unknown funtion
C45E5.4	unknown funtion
C46A5.6	unknown funtion
C55B7.1	glh-2: Germ-Line Helicase; ATP-dependent RNA helicase ;
C57115_rc / yk280a11.3	Y54G2A.9: C-type lectin
C57244 / yk301a11.3	rps-13: 40S ribosomal protein S13;
D1086.1	DUF19
D1086.3	DUF19
D2030.7	Zn-finger, C2H2 type
F07G6.7	FTH domain & DUF38
F11A5.3	GTPase Rab2, small G protein superfamily, Ras family
F13D12.5	unknown funtion
F14F8.6	srw-44: Serpentine Receptor, class W; 7-transmembrane olfactory receptor
F19C7.2	Hydrolytic enzymes of the alpha/beta hydrolase fold
F20A1.9	Putative DEAD-box RNA helicase DDX1
F37B1.4	gst-15: Glutathione S-Transferase
F41E6.14	Integral membrane O-acyltransferase
F44C8.1	cyp-33C4: Cytochrome P450 CYP2 subfamily
F44F1.6B	Uncharacterized coiled-coil containing protein
F46F11.2	cey-2: C.Elegan cold-shock/Y-box protein; Predicted RNA-binding protein containing PIN domain and invovled in translation or RNA processing;

F49E12.1	Peroxidase/oxygenase
F53G12.8	unknown funtion
F56C9.4	unknown funtion
F59A7.2	unknown funtion
K08B12.2	Transcription factor Doublesex
K11G9.5	Permease of the major facilitator superfamily
M02D8.4	Asparagine synthase (glutamine-hydrolyzing)
M7.2	klc-1: Kinesin light chain
R07C12.1	R07C12.1 : 7-transmembrane receptor
R09B5.3	cnc-2: Caenacin (Caenorhabditis bacteriocin)
R144.6	unknown funtion
T05A7.3	C-type lectin
T07A5.1	domain of Fukutin-related
T07G12.11	domains of Short-chain dehydrogenase/reductase and Zn-finger, C2H2 type
T12B3.2	Synaptic vesicle transporter SVOP and related transporters (major facilitator superfamily)
T13B5.5	Triacylglycerol lipase
T20F7.3	Inositol polyphosphate multikinase, component of the ARGR transcription regulatory complex
T21B10.4	unknown funtion
T21F4.1	Arginase; orthologous to the human gene Arginase type I erythroid variant
T22C8.5	Zn-finger C2H2 type
T23F4.4	nas-27: Nematode Astacin protease; Meprin A metalloprotease
T24C4.1	Ubiquinol cytochrome c reductase, subunit QCR2
T26E4.12	srd-27: Serpentine Receptor, class D (delta); Chemoreceptor/7TM receptor
T28D6.3	unknown funtion
W02A2.6	rec-8: a meiosis-specific cohesin complex subunit; REC-8 is essential for pairing of homologoues and sister chromatids during meiosis and thus for proper chromosome disjunction;
W05F2.6	unknown funtion
W06A7.4	unknown funtion
Y110A7A.15	unknown funtion
Y111B2A.1	Lammer dual specificity kinases
Y37A1A.2	DUF895 & Major Facilitator Superfamily domain
Y37H2A.4	containing an F-box and an FTH/DUF38 motif, both mediate protein-protein interaction.
Y38C1AB.2	containing domain of Cation (not K+) channel, TM region
Y38E10A.13	DUF1412
Y59A8B.11	contains FTH domain and F-box domain
Y69A2AR.V	Y69A2AR.13
Y75B12B.1	unknown funtion
yk560b8.3	Y54G2A.8:C-type lectin
yk577g10.5	T22F3.11 : Permease of the major facilitator superfamily
yk620h12.3	C25A1.8: C-type lectin
yk625b7.3	C31H1.8
yk724g12.3	Y38E10A.14
yk727b8.3	T26F2.2: Nuclear 5'-3' exoribonuclease-interacting protein, Rai1p

ZK1058.6	Carbon-nitrogen hydrolase
ZK418.8	Vigilin
ZK488.7	pqn-98: Prion-like-(Q/N-rich)-domain-bearing protein; Predicted alpha-helical protein;
ZK6.8	DUF895
ZK666.6	Lectin C-type domain/CUB domain

Genes downregulated in *sma-9* background

6R55.1	Beta-2-glycoprotein I
B0280.7	unknown function
C54C8.4	Glycosyltranferase ; DUF273
F01E11.5	transmembrane receptor ; rhodopsin family
F02C12.5C	cyp-13B1: Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies
F15E11.14	unknown function
F16F9.2	dpy-6: Dumpy gene; shorter than wild-type
F17E9.11	lys-10: N-acetylmuraminidase/lysozyme
F23F12.9	Basic region leucine zipper transcription factor
F54F11.2	M13 family peptidase; encodes a neprilysin; neprilysins are thermolysin-like zinc metallopeptidases,
F55G11.2	DUF141
F55G11.4	DUF141
K02E10.4	Worm-specific repeat type 1 doamin
K10D11.1	dod-17: Downstream Of DAF-16 (regulated by DAF-16)
M02F4.7	C-type lectin
M03B6.3	unknown function
T16G12.1	Puromycin-sensitive aminopeptidase and related aminopeptidases
Y116F11B.3	pcp-4: Prolyl Carboxy Peptidase like; Hydrolytic enzymes of the alpha/beta hydrolase fold
Y51H4A.5	Predicted lipase
ZC449.3	Mitogen-activated protein kinase (MAPK) kinase MKK4
ZK287.4	Serine proteinase inhibitor (KU family)
ZK896.7	C-type lectin

Table 8. Genes regulated by *dbl-1* and *sma-9* in DNA microarray analysis

Genes upregulated in *dbl-1* and *sma-9* background

Probe194153_at	multi target
C01A2.4	DUF279
C04E12.2	DUF19
C04E12.5	DUF750
C08E3.1	unknown function
C17B7.2	DUF19
C17H12.6	DUF141
C18E3.8	hop-1: Homolog Of Presenilin; Peptidase A22, presenilin signal peptide
C30G12.6	unknown function
C36C5.5	DUF19
C38C10.2	Permease of the major facilitator superfamily ; human gene Solute carrier family 17 (Anion/Sugar transporter)
F01G10.3	ech-9: Enoyl-CoA Hydratase gene class
F14H3.12	Serine/threonine protein kinase
F15B9.1	far-3: Nematode fatty acid retinoid binding protein (Gp-FAR-1)
F16H6.7	Predicted pyroglutamyl peptidase; DUF976
F19B2.3	srw-39: Serpentine Receptor, class W; 7-transmembrane olfactory receptor
F21F8.4	aspartyl protease
F31F4.15	FTH domain; DUF38
F35E2.9	DUF316//Peptidase
F35F10.6	DUF976
F49C12.7	DUF227
F55G1.5	Mitochondrial solute carrier protein
F57G4.5	unknown function
F57G4.6	MADF domain- Interpro domain
H36L18.2	similarity to Ashbya gossypii
K01D12.14	cdr-5: Cadmium Responsive gene; Glutathione S-transferases/Failed axon connections (fax) protein
K07A1.1	unknown function
K07A1.4	similar to African swine fever virus Major capsid protein
K09G1.2	unknown function
M01E11.5	cey-3: C.Elegans cold-shock/Y-box protein; Predicted RNA-binding protein containing PIN domain and invovled in translation or RNA processing
R09B5.6	3-hydroxyacyl-CoA dehydrogenase
T05G5.10	iff-1: Initiation Factor Five (eIf-5A) homologs required for germ cell proliferation
T06C12.4	FTH domain // DUF38
T06E6.10	Fibrillins and related proteins containing Ca2+-binding EGF-like domains
T20D4.11	DUF19
T20D4.12	DUF19
T21G5.3	glh-1: Germ-Line Helicase; DEAD-box RNA helicase; germ-line development
T27F2.4	Basic-leucine zipper (bZIP) transcription factor domain
T28A11.16	DUF19
T28A11.2	DUF19

W08E12.2	domain of 2Fe-2S ferredoxin, iron-sulfur binding site
Y111B2A.2	unknown function
Y38E10A.15	DUF1412
Y46C8AL.D	Y46C8AL.3: type C lectin
Y47D7A.12	similarity to Homo sapiens RB protein binding protein; RB protein(Retinoblastoma tumor suppressor gene) binding protein
Y69H2.9	unknown function
yk220a12.3	yk220a12.3
yk325c3.3	Y73B6BL.23
yk391c10.3	C18E3.5: U5 snRNP-specific protein-like factor and related proteins
yk484h4.3	gld-1: defective in Germ Line Development; K homology RNA binding domain required for meiotic cell cycle progression
yk566h11.3	B0280.5: a direct target of GLD-1; a protein with a chitin binding peritrophin-A domain that is required, in conjunction with CEJ-1, for early embryonic development;
yk641c8.5	K10B3.6: Predicted starch-binding protein
yk716c10.3	Y17D7B.7
yk93e2.3	mex-5: Muscle Excess gene; CCCH-type Zn-finger protein
ZK1251.2	ins-7: insulin-like peptide; is one of 38 insulin-like peptides in C. elegans
ZK250.1	Endonuclease/exonuclease/phosphatase //RNA-directed DNA polymerase (Reverse transcriptase)
ZK381.1	him-3: High Incidence of Males (increased X chromosome loss)
ZK669.3	Gamma interferon inducible lysosomal thiol reductase
ZK970.7	DUF 148

Genes downregulated in *dbl-1* and *sma-9* background

B0218.8	Lectin C-type
C24B9.3	von Willebrand factor type A domain
C47D12.1	trr-1: TRRAP-like (transcription/transformation domain-associated protein); Histone acetyltransferase SAGA, TRRAP/TRA1 component, PI-3 kinase superfamily
F09G2.3	Na+/Pi symporter ; Phosphate transporter family
F10D2.9	fat-7: an essential delta-9 fatty acid desaturase required for the synthesis of monounsaturated fatty acids
F15E11.12	unknown function
F35C5.9	Lectin C-type domain/ CUB domain
F49E10.2	Beta, beta-carotene 15,15'-dioxygenase and related enzymes ; domain of Retinal pigment epithelial membrane protein
F56A4.2	Lectin C-type
K08E7.5	unknown function
K10C2.3	aspartyl proteases
T13F2.8	cav-1: caveolin protein family
W01F3.3	Serine proteinase inhibitor (KU family)
W03G1.7	asm-3: Acid Sphingomyelinase
Y19D10B.A	unknown function
Y38E10A.5	Type C lectin
yk716h9.3	F15E11.15
Y71D11A.1	contains cadherin repeats

ZK617.2	Triacylglycerol lipase
ZK678.5	wrt-4: Warthog (hedgehog-like family); Sonic hedgehog

REFERENCES:

Affolter M, Marty T, Vigano MA, Jazwinska A. (2001) Nuclear interpretation of Dpp signaling in Drosophila. EMBO J. 20(13):3298-305.

Arduengo PM, Appleberry OK, Chuang P, L'Hernault SW. (1998) The presenilin protein family member SPE-4 localizes to an ER/Golgi derived organelle and is required for proper cytoplasmic partitioning during Caenorhabditis elegans spermatogenesis. J Cell Sci. 111 (Pt 24):3645-54.

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.

Aspock G, Kagoshima H, Niklaus G, Burglin TR. (1999) Caenorhabditis elegans has scores of hedgehog-related genes: sequence and expression analysis. Genome Res. 9(10):909-23.

Bailis JM, Forsburg SL. (2004) MCM proteins: DNA damage, mutagenesis and repair. Curr Opin Genet Dev. 14(1):17-21.

Baird SE, Ellazar SA. (1999) TGFbeta-like signaling and spicule development in Caenorhabditis elegans. Dev Biol. 212(1):93-100.

Baker JC, Beddington RS, Harland RM. (1999) Wnt signaling in Xenopus embryos inhibits bmp4 expression and activates neural development. Genes Dev. 13(23):3149-59.

Balemans W, Van Hul W. (2002) Extracellular regulation of BMP signaling in vertebrates: a cocktail of modulators. Dev Biol. 250(2):231-50.

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.

Blokzijl A, Dahlqvist C, Reissmann E, Falk A, Moliner A, Lendahl U, Ibanez CF. (2003) Cross-talk between the Notch and TGF-beta signaling pathways mediated by interaction of the Notch intracellular domain with Smad3. J Cell Biol. 163(4):723-8.

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, 851-4.

Bonni S, Wang HR, Causing CG, Kavsak P, Stroschein SL, Luo K, Wrana JL. (2001) TGF-beta induces assembly of a Smad2-Smurf2 ubiquitin ligase complex that targets SnoN for degradation. Nat Cell Biol. 3(6):587-95.

Bower KE, Zeller RW, Wachsman W, Martinez T, McGuire KL. (2002) Correlation of transcriptional repression by p21(SNFT) with changes in DNA.NF-AT complex interactions. J Biol Chem. 277(38):34967-77.

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, 553-62.

Capdevila J, Izpisua Belmonte JC. (2001) Patterning mechanisms controlling vertebrate limb development. Annu Rev Cell Dev Biol. 17:87-132.

C. elegans Sequencing Consortium. (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. Science. 282(5396):2012-8.

Chamberlin HM, Thomas JH. (2000) The bromodomain protein LIN-49 and trithorax-related protein LIN-59 affect development and gene expression in Caenorhabditis elegans. Development. 127(4):713-23.

Cheifetz S, Weatherbee JA, Tsang ML, Anderson JK, Mole JE, Lucas R, Massague J. (1987) The transforming growth factor-beta system, a complex pattern of cross-reactive ligands and receptors. Cell. 48(3):409-15

Chen W, Kirkbride KC, How T, Nelson CD, Mo J, Frederick JP, Wang XF, Lefkowitz RJ, Blobe GC. (2003) Beta-arrestin 2 mediates endocytosis of type III TGF-beta receptor and down-regulation of its signaling. Science. 301(5638):1394-7

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

Colavita A, Krishna S, Zheng H, Padgett RW, Culotti JG (1998) Pioneer axon guidance by UNC-129, a C. elegans TGF-beta. Science. 281(5377):706-9.

Collet, J., Spike, C. A., Lundquist, E. A., Shaw, J. E., 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.

Cordenonsi M, Dupont S, Maretto S, Insinga A, Imbriano C, Piccolo S (2003) Links between tumor suppressors: p53 is required for TGF-beta gene responses by cooperating with Smads. Cell. 113(3):301-14.

Couillault C, Pujol N, Reboul J, Sabatier L, Guichou JF, Kohara Y, Ewbank JJ. (2004) TLR-independent control of innate immunity in Caenorhabditis elegans by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. Nat Immunol. 5(5):488-94

da Graca LS, Zimmerman KK, Mitchell MC, Kozhan-Gorodetska M, Sekiewicz K, Morales Y, Patterson GI. (2004) DAF-5 is a Ski oncoprotein homolog that functions

in a neuronal TGF beta pathway to regulate C. elegans dauer development. Development. 131(2):435-46.

Dahlqvist C, Blokzijl A, Chapman G, Falk A, Dannaeus K, Ibanez CF, Lendahl U. (2003) Functional Notch signaling is required for BMP4-induced inhibition of myogenic differentiation. Development. 130(24):6089-99.

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.

Deacu E, Mori Y, Sato F, Yin J, Olaru A, Sterian A, Xu Y, Wang S, Schulmann K, Berki A, Kan T, Abraham JM, Meltzer SJ. (2004) Activin type II receptor restoration in ACVR2-deficient colon cancer cells induces transforming growth factor-beta response pathway genes. Cancer Res. 64(21):7690-6.

Dennler S, Itoh S, Vivien D, ten Dijke P, Huet S, Gauthier JM. (1998) Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. EMBO J. 17(11):3091-100.

Derynck R, Feng XH. (1997) TGF-beta receptor signaling. Biochim Biophys Acta. 1333(2):F105-50.

Derynck R, Jarrett JA, Chen EY, Eaton DH, Bell JR, Assoian RK, Roberts AB, Sporn MB, Goeddel DV. (1985) Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells. Nature. 316(6030):701-5

Derynck R, Zhang YE. (2003) Smad-dependent and Smad-independent pathways in TGF-beta family signalling. Nature. 425(6958):577-84.

Derynck R, Zhang Y, Feng XH. (1998) Smads: transcriptional activators of TGF-beta responses. Cell. 95(6):737-40.

Di Guglielmo GM, Le Roy C, Goodfellow AF, Wrana JL. (2003) Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. Nat Cell Biol. 5(5):410-21.

Ebisawa T, Fukuchi M, Murakami G, Chiba T, Tanaka K, Imamura T, Miyazono K.(2001) Smurf1 interacts with transforming growth factor-beta type I receptor through Smad7 and induces receptor degradation. J Biol Chem. 276(16):12477-80.

Eckmann CR, Kraemer B, Wickens M, Kimble J. (2002) GLD-3, a bicaudal-C homolog that inhibits FBF to control germline sex determination in C. elegans. Dev Cell. 3(5):697-710.

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.

Flemming AJ, Shen ZZ, Cunha A, Emmons SW, Leroi AM. (2000) Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. Proc Natl Acad Sci U S A.: 97(10):5285-90.

Fukuchi M, Imamura T, Chiba T, Ebisawa T, Kawabata M, Tanaka K, Miyazono K. (2001) Ligand-dependent degradation of Smad3 by a ubiquitin ligase complex of ROC1 and associated proteins. Mol Biol Cell. 12(5):1431-43.

Garcia LR, Mehta P, Sternberg PW. (2001) Regulation of distinct muscle behaviors controls the C. elegans male's copulatory spicules during mating. Cell. 107(6):777-88.

Gascoigne, N. R. (2001). Positive selection in a Schnurri. Nat Immunol. 2, 989-91.

Georgi LL, Albert PS, Riddle DL. (1990) daf-1, a C. elegans gene controlling dauer larva development, encodes a novel receptor protein kinase. Cell. 61(4):635-45.

Gerisch B, Weitzel C, Kober-Eisermann C, Rottiers V, Antebi A. (2001) A hormonal signaling pathway influencing C. elegans metabolism, reproductive development, and life span. Dev Cell. 1(6):841-51

Germain S, Howell M, Esslemont GM, Hill CS. (2000) Homeodomain and winged-helix transcription factors recruit activated Smads to distinct promoter elements via a common Smad interaction motif. Genes Dev. 14(4):435-51

Goumans MJ, Valdimarsdottir G, Itoh S, Rosendahl A, Sideras P, ten Dijke P. (2002) Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors. EMBO J. 21(7):1743-53.

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.

Grimm OH, Gurdon JB. (2002) Nuclear exclusion of Smad2 is a mechanism leading to loss of competence. Nat Cell Biol. 4(7):519-22.

Gumienny TL, Padgett RW (2003) A small issue addressed. Bioessays. 25(4):305-8.

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, 3337-47.

Hansen D, Hubbard EJ, Schedl T. (2004) Multi-pathway control of the proliferation versus meiotic development decision in the Caenorhabditis elegans germline. Dev Biol. 268(2):342-57.

Hansen D, Wilson-Berry L, Dang T, Schedl T. (2004) Control of the proliferation versus meiotic development decision in the C. elegans germline through regulation of GLD-1 protein accumulation. Development. 131(1):93-104.

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, 229-40.

Hayes S, Chawla A, Corvera S. (2002) TGF beta receptor internalization into EEA1-enriched early endosomes: role in signaling to Smad2. J Cell Biol. 158(7):1239-49.

Hemler ME. (2003) Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain. Annu Rev Cell Dev Biol. 19:397-422.

Higgins CF, Linton KJ. (2004) The ATP switch model for ABC transporters. Nat Struct Mol Biol. 11(10):918-26.

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.

Hirose T, Nakano Y, Nagamatsu Y, Misumi T, Ohta H, Ohshima Y. (2003) Cyclic GMP-dependent protein kinase EGL-4 controls body size and lifespan in C elegans. Development. 130(6):1089-99.

Hobert O. (2002) PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. Biotechniques. 32(4):728-30.

Hogan, B. L. (1996). Bone morphogenetic proteins in development. Curr Opin Genet Dev. 6, 432-8.

Hsu AL, Murphy CT, Kenyon C. (2003) Regulation of aging and age-related disease by DAF-16 and heat-shock factor. Science. 300(5622):1142-5.

Hua X, Miller ZA, Wu G, Shi Y, Lodish HF. (1999) Specificity in transforming growth factor beta-induced transcription of the plasminogen activator inhibitor-1 gene: interactions of promoter DNA, transcription factor muE3, and Smad proteins. Proc Natl Acad Sci U S A. 96(23):13130-5.

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.

Huse M, Chen YG, Massague J, Kuriyan J. (1999) Crystal structure of the cytoplasmic domain of the type I TGF beta receptor in complex with FKBP12. Cell. 96(3):425-36.

Ito Y, Miyazono K. (2003) RUNX transcription factors as key targets of TGF-beta superfamily signaling. Curr Opin Genet Dev. 13(1):43-7.

Itoh S, Itoh F, Goumans MJ, Ten Dijke P. (2000) Signaling of transforming growth factor-beta family members through Smad proteins. Eur J Biochem. 267(24):6954-67.

Izzi L, Attisano L. (2004) Regulation of the TGFbeta signalling pathway by ubiquitin-mediated degradation. Oncogene. 23(11):2071-8.

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, 563-73.

Ji YJ, Nam S, Jin YH, Cha EJ, Lee KS, Choi KY, Song HO, Lee J, Bae SC, Ahnn J. (2004) RNT-1, the C. elegans homologue of mammalian RUNX transcription factors, regulates body size and male tail development. Dev Biol. 274(2):402-12.

Kamath RS, Ahringer J. (2003) Genome-wide RNAi screening in Caenorhabditis elegans. Methods. 30(4):313-21.

Kavsak P, Rasmussen RK, Causing CG, Bonni S, Zhu H, Thomsen GH, Wrana JL.(2000) Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation. Mol Cell. 6(6):1365-75.

Kawase E, Wong MD, Ding BC, Xie T. (2004) Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing bam transcription in the Drosophila testis. Development. 131(6):1365-75.

Kimble J, Hirsh D. (1979) The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. Dev Biol. 70(2):396-417.

Kramer JM, French RP, Park EC, Johnson JJ. (1990) The Caenorhabditis elegans rol-6 gene, which interacts with the sqt-1 collagen gene to determine organismal morphology, encodes a collagen. Mol Cell Biol. 10(5):2081-9.

Kretzschmar M, Liu F, Hata A, Doody J, Massague J (1997) The TGF-beta family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. Genes Dev. 11(8):984-95.

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, 251-60.

Kuchin S, Carlson M. (2003). Analysis of transcriptional repression by Mig1 in Saccharomyces cerevisiae using a reporter assay. Methods Enzymol.371:602-14.

Kurokawa M, Mitani K, Imai Y, Ogawa S, Yazaki Y, Hirai H (1998) The t(3;21) fusion product, AML1/Evi-1, interacts with Smad3 and blocks transforming growth factor-beta-mediated growth inhibition of myeloid cells. Blood. 92(11):4003-12.

Kurz CL, Tan MW. (2004) Regulation of aging and innate immunity in C. elegans. Aging Cell. 3(4):185-93.

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, 153-8.

Lambie EJ. (2002) Cell proliferation and growth in C. elegans. Bioessays. 24(1):38-53.

Lamont LB, Crittenden SL, Bernstein D, Wickens M, Kimble J. (2004) FBF-1 and FBF-2 regulate the size of the mitotic region in the C. elegans germline. Dev Cell. 7(5):697-707.

Lee MH, Schedl T. (2001) Identification of in vivo mRNA targets of GLD-1, a maxi-KH motif containing protein required for C. elegans germ cell development. Genes Dev. 15(18):2408-20.

Levy S, Shoham T. (2005) The tetraspanin web modulates immune-signalling complexes. Nat Rev Immunol. 5(2):136-48.

Liang J, Lints R, Foehr ML, Tokarz R, Yu L, Emmons SW, Liu J, Savage-Dunn C. (2003) The Caenorhabditis elegans schnurri homolog sma-9 mediates stage- and cell type-specific responses to DBL-1 BMP-related signaling. Development. 130(26):6453-64.

Li X, Greenwald I. (1997) HOP-1, a Caenorhabditis elegans presenilin, appears to be functionally redundant with SEL-12 presenilin and to facilitate LIN-12 and GLP-1 signaling. Proc Natl Acad Sci U S A. 94(22):12204-9.

Li Y, Turck CM, Teumer JK, Stavnezer E (1986) Unique sequence, ski, in Sloan-Kettering avian retroviruses with properties of a new cell-derived oncogene. J Virol. 57(3):1065-72.

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, 5819-31.

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, 2390-402.

Liu F, Hata A, Baker JC, Doody J, Carcamo J, Harland RM, Massague J. (1996) A human Mad protein acting as a BMP-regulated transcriptional activator. Nature. 381(6583):620-3.

Liu T, Zimmerman KK, Patterson GI. (2004) Regulation of signaling genes by TGFbeta during entry into dauer diapause in C. elegans. BMC Dev Biol. 4(1):11.

Li Y, Turck CM, Teumer JK, Stavnezer E. (1986) Unique sequence, ski, in Sloan-Kettering avian retroviruses with properties of a new cell-derived oncogene. J Virol. 57(3):1065-72.

Liu X, Sun Y, Weinberg RA, Lodish HF. (2001) Ski/Sno and TGF-beta signaling. Cytokine Growth Factor Rev. 12(1):1-8.

Lo RS, Wotton D, Massague J. (2001) Epidermal growth factor signaling via Ras controls the Smad transcriptional co-repressor TGIF. EMBO J. 20(1-2):128-36.

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.

Loris R. (2002) Principles of structures of animal and plant lectins. Biochim Biophys Acta. 1572(2-3):198-208.

Luo K, Lodish HF (1997) Positive and negative regulation of type II TGF-beta receptor signal transduction by autophosphorylation on multiple serine residues. EMBO J. 16(8):1970-81.

Luo K. (2004) Ski and SnoN: negative regulators of TGF-beta signaling. Curr Opin Genet Dev. 14(1):65-70.

Macias-Silva M, Abdollah S, Hoodless PA, Pirone R, Attisano L, Wrana JL. (1996) MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. Cell. 87(7):1215-24.

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, 418-28.

Maduzia LL, Roberts AF, Wang H, Lin X, Chin L, Zimmerman CM, Cohen S, Feng XH, Padgett RW. (2005) C. elegans serine-threonine kinase KIN-29 modulates TGFb signaling and regulates body size formation. BMC Dev Biol. 5(1):8

Maeda I, Kohara Y, Yamamoto M, Sugimoto A. (2001) Large-scale analysis of gene function in Caenorhabditis elegans by high-throughput RNAi. Curr Biol. 11(3):171-6.

Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, Kohara Y, Ewbank JJ. (2002) Inducible antibacterial defense system in C. elegans. Curr Biol. 12(14):1209-14.

Marty, T., Muller, B., Basler, K. and Affolter, M. (2000). Schnurri mediates Dpp-dependent repression of brinker transcription. Nat Cell Biol. 2, 745-9.

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

Massague J. (2000) How cells read TGF-beta signals. Nat Rev Mol Cell Biol. 1(3):169-78.

Massague J, Chen YG. (2000) Controlling TGF-beta signaling. Genes Dev. 14(6):627-44.

Massague J, Wotton D. (2000) Transcriptional control by the TGF-beta/Smad signaling system. EMBO J. 19(8):1745-54.

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.

Miller MA, Cutter AD, Yamamoto I, Ward S, Greenstein D. (2004) Clustered organization of reproductive genes in the C. elegans genome. Curr Biol;14(14):1284-90.

Millet AC, Ewbank JJ. (2004) Immunity in Caenorhabditis elegans. Curr Opin Immunol. 16(1):4-9.

Miskowski J, Li Y, Kimble J. (2001) The sys-1 gene and sexual dimorphism during gonadogenesis in Caenorhabditis elegans. Dev Biol. 230(1):61-73.

Mochii M, Yoshida S, Morita K, Kohara Y, Ueno N. (1999) Identification of transforming growth factor-beta- regulated genes in caenorhabditis elegans by differential hybridization of arrayed cDNAs. Proc Natl Acad Sci U S A. 96(26):15020-5.

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, 1337-47.

Morita K, Shimizu M, Shibuya H, Ueno N. (2001) A DAF-1-binding protein BRA-1 is a negative regulator of DAF-7 TGF-beta signaling. Proc Natl Acad Sci U S A. 98(11):6284-8. Epub 2001 May 15.

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, 1063-73.

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.

Murphy CT, McCarroll SA, Bargmann CI, Fraser A, Kamath RS, Ahringer J, Li H, Kenyon C. (2003) Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature. 424(6946):277-83.

Myllyharju J, Kivirikko KI. (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet. 2004 Jan;20(1):33-43.

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., Christian, J. L. (2000). Regulation of BMP/Dpp signaling during embryonic development. Cell Mol Life Sci. 57, 943-56.

Nakashima K, Yanagisawa M, Arakawa H, Kimura N, Hisatsune T, Kawabata M, Miyazono K, Taga T. (1999) Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300. Science. 284(5413):479-82.

Nam S, Jin YH, Li QL, Lee KY, Jeong GB, Ito Y, Lee J, Bae SC. (2002) Expression pattern, regulation, and biological role of runt domain transcription factor, run, in Caenorhabditis elegans. Mol Cell Biol. 22(2):547-54.

Newfeld SJ, Wisotzkey RG, Kumar S. (1999) Molecular evolution of a developmental pathway: phylogenetic analyses of transforming growth factor-beta family ligands, receptors and Smad signal transducers. Genetics. 152(2):783-95.

Nicholas HR, Hodgkin J. (2004) Responses to infection and possible recognition strategies in the innate immune system of Caenorhabditis elegans. Mol Immunol. 41(5):479-93.

Nishita M, Hashimoto MK, Ogata S, Laurent MN, Ueno N, Shibuya H, Cho KW. (2000) Interaction between Wnt and TGF-beta signalling pathways during formation of Spemann's organizer. Nature. 403(6771):781-5.

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.

Oka Y, Omura M, Kataoka H, Touhara K. (2004) Olfactory receptor antagonism between odorants. EMBO J. 23(1):120-6.

Oklu R, Hesketh R. (2000) The latent transforming growth factor beta binding protein (LTBP) family. Biochem J. 352 Pt 3:601-10.

Ookuma S, Fukuda M, Nishida E. (2003) Identification of a DAF-16 transcriptional target gene, scl-1, that regulates longevity and stress resistance in Caenorhabditis elegans. Curr Biol. 13(5):427-31.

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, 121-31.

Patel MN, Knight CG, Karageorgi C, Leroi AM. (2002) Evolution of germ-line signals that regulate growth and aging in nematodes. Proc Natl Acad Sci U S A. 99(2):769-74.

Patterson GI, Koweek A, Wong A, Liu Y, Ruvkun G. (1997) The DAF-3 Smad protein antagonizes TGF-beta-related receptor signaling in the Caenorhabditis elegans dauer pathway. Genes Dev. 11(20):2679-90.

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

Pyrowolakis G, Hartmann B, Muller B, Basler K, Affolter M. (2004) A simple molecular complex mediates widespread BMP-induced repression during Drosophila development. Dev Cell. 7(2):229-40.

Raftery LA, Twombly V, Wharton K, Gelbart WM. (1995) Genetic screens to identify elements of the decapentaplegic signaling pathway in Drosophila. Genetics. 139(1):241-54.

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, 145-56.

Remacle JE, Kraft H, Lerchner W, Wuytens G, Collart C, Verschueren K, Smith JC, Huylebroeck D. (1999) New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. EMBO J. 18(18):5073-84.

Ren P, Lim CS, Johnsen R, Albert PS, Pilgrim D, Riddle DL(1996) Control of C. elegans larval development by neuronal expression of a TGF-beta homolog. Science. 274(5291):1389-91.

Roussa E, Krieglstein K. (2004) Induction and specification of midbrain dopaminergic cells: focus on SHH, FGF8, and TGF-beta. Cell Tissue Res. 318(1):23-33

Rushlow C, Colosimo PF, Lin MC, Xu M, Kirov N. (2001) Transcriptional regulation of the Drosophila gene zen by competing Smad and Brinker inputs. Genes Dev. 15(3):340-51.

Ryder SP, Frater LA, Abramovitz DL, Goodwin EB, Williamson JR. (2004) RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1. Nat Struct Mol Biol. 11(1):20-8.

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 U S A. 93, 790-4.

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

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, 239-47.

Scherer A, Graff JM. (2000) Calmodulin differentially modulates Smad1 and Smad2 signaling. J Biol Chem. 275(52):41430-8.

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.

Sekelsky JJ, Newfeld SJ, Raftery LA, Chartoff EH, Gelbart WM. (1995) Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in Drosophila melanogaster. Genetics. 139(3):1347-58.

Shi X, Yang X, Chen D, Chang Z, Cao X(1999) Smad1 interacts with homeobox DNA-binding proteins in bone morphogenetic protein signaling. J Biol Chem. 274(19):13711-7.

Shi Y, Massague J. (2003) Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell. 113(6):685-700.

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, 585-94.

Shostak Y, Van Gilst MR, Antebi A, Yamamoto KR. (2004) Identification of C. elegans DAF-12-binding sites, response elements, and target genes. Genes Dev. 18(20):2529-44Silver SJ, Rebay I. (2005) Signaling circuitries in development: insights from the retinal determination gene network. Development. 132(1):3-13.

Silver SJ, Rebay I. (2005) Signaling circuitries in development: insights from the retinal determination gene network. Development. 132(1):3-13.

Sivasankaran R, Vigano MA, Muller B, Affolter M, Basler K. (2000) Direct transcriptional control of the Dpp target omb by the DNA binding protein Brinker. EMBO J. 19(22):6162-72.

Song X, Wong MD, Kawase E, Xi R, Ding BC, McCarthy JJ, Xie T. (2004) Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. Development. 131(6):1353-64.

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, 3393-403.

Steneberg P, Hemphala J, Samakovlis C. (1999) Dpp and Notch specify the fusion cell fate in the dorsal branches of the Drosophila trachea. Mech Dev. 87(1-2):153-63.

Stroschein SL, Wang W, Zhou S, Zhou Q, Luo K. (1999) Negative feedback regulation of TGF-beta signaling by the SnoN oncoprotein. Science. 286(5440):771-4.

Sugawara K, Morita K, Ueno N, Shibuya H. (2001) BIP, a BRAM-interacting protein involved in TGF-beta signalling, regulates body length in Caenorhabditis elegans. Genes Cells. 2001 Jul;6(7):599-606.

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

Sulston JE, Albertson DG, Thomson JN. (1980) The Caenorhabditis elegans male: postembryonic development of nongonadal structures. Dev Biol. 78(2):542-76.

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, 1631-9.

Sze JY, Liu Y, Ruvkun G. (1997) VP16-activation of the C. elegans neural specification transcription factor UNC-86 suppresses mutations in downstream genes and causes defects in neural migration and axon outgrowth. Development. 124(6):1159-68.

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.

Tabata T, Takei Y. (2004) Morphogens, their identification and regulation. Development:131 (4):703-12.

Takagi, T., Harada, J. and Ishii, S. (2001). Murine Schnurri-2 is required for positive selection of thymocytes. Nat Immunol. 2, 1048-53.

Tan MW, Ausubel FM. (2000) Caenorhabditis elegans: a model genetic host to study Pseudomonas aeruginosa pathogenesis. Curr Opin Microbiol. 3(1):29-34.

Tan MW (2001) Genetic and genomic dissection of host-pathogen interactions using a *P. aeruginosa- C. elegans* pathogenesis model. Pediatric Pulmonol. 32, 96-97.

ten Dijke P, Hill CS. (2004) New insights into TGF-beta-Smad signalling. Trends Biochem Sci. 29(5):265-73.

Tolkunova EN, Fujioka M, Kobayashi M, Deka D, Jaynes JB. (1998) Two distinct types of repression domain in engrailed: one interacts with the groucho corepressor and is preferentially active on integrated target genes. Mol Cell Biol. 18(5):2804-14.

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, 1657-70.

Troemel ER, Chou JH, Dwyer ND, Colbert HA, Bargmann CI. (1995) Divergent seven transmembrane receptors are candidate chemosensory receptors in C. elegans. Cell. 83(2):207-18.

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 U S A. 89, 8971-5.

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.

Wan Y, Liu X, Kirschner MW. (2001) The anaphase-promoting complex mediates TGF-beta signaling by targeting SnoN for destruction. Mol Cell. 8(5):1027-39.

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, 4989-98.

Ward S, Burke DJ, Sulston JE, Coulson AR, Albertson DG, Ammons D, Klass M, Hogan E. (1988) Genomic organization of major sperm protein genes and pseudogenes in the nematode Caenorhabditis elegans. J Mol Biol.199(1):1-13.

Watanabe M, Whitman M. (1999) FAST-1 is a key maternal effector of mesoderm inducers in the early Xenopus embryo. Development. 126(24):5621-34.

Wells RG, Yankelev H, Lin HY, Lodish HF. (1997) Biosynthesis of the type I and type II TGF-beta receptors. Implications for complex formation. J Biol Chem. 272(17): 11444-51.

Wotton D, Lo RS, Lee S, Massague J. (1999) A Smad transcriptional corepressor. Cell. 97(1):29-39.

Wotton D, Massague J. (2001) Smad transcriptional corepressors in TGF beta family signaling. Curr Top Microbiol Immunol. 254:145-64.

Wrana JL, Attisano L, Wieser R, Ventura F, Massague J. (1994) Mechanism of activation of the TGF-beta receptor. Nature. 370(6488):341-7.

Wu JW, Krawitz AR, Chai J, Li W, Zhang F, Luo K, Shi Y. (2002) Structural mechanism of Smad4 recognition by the nuclear oncoprotein Ski: insights on Ski-mediated repression of TGF-beta signaling. Cell. 111(3):357-67.

Xu H, Staszewski L, Tang H, Adler E, Zoller M, Li X. (2004) Different functional roles of T1R subunits in the heteromeric taste receptors. Proc Natl Acad Sci U S A. 101(39):14258-63.

Xu W, Angelis K, Danielpour D, Haddad MM, Bischof O, Campisi J, Stavnezer E, Medrano EE. (2000) Ski acts as a co-repressor with Smad2 and Smad3 to regulate the response to type beta transforming growth factor. Proc Natl Acad Sci U S A. 97(11):5924-9.

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.

Zawel L, Dai JL, Buckhaults P, Zhou S, Kinzler KW, Vogelstein B, Kern SE. (1998) Human Smad3 and Smad4 are sequence-specific transcription activators. Mol Cell. 1(4):611-7.

Zhang Y, Feng XH, Derynck R. (1998) Smad3 and Smad4 cooperate with c-Jun/c-Fos to mediate TGF-beta-induced transcription. Nature. 394(6696):909-13.

Zhang YW, Yasui N, Ito K, Huang G, Fujii M, Hanai J, Nogami H, Ochi T, Miyazono K, Ito Y. (2000) A RUNX2/PEBP2alpha A/CBFA1 mutation displaying impaired transactivation and Smad interaction in cleidocranial dysplasia. Proc Natl Acad Sci U S A. 97(19):10549-54.

Zhou S, Zawel L, Lengauer C, Kinzler KW, Vogelstein B. (1998) Characterization of human FAST-1, a TGF beta and activin signal transducer. Mol Cell. 2(1):121-7.

Zhu H, Kavsak P, Abdollah S, Wrana JL, Thomsen GH. (1999) A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. Nature. 400(6745):687-93.

Zimmerman CM, Kariapper MS, Mathews LS. (1998) Smad proteins physically interact with calmodulin. J Biol Chem. 273(2):677-80.