DBL-1/BMP target genes and *sma-9* function in *C. elegans*

By

Jianghua Yin

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

2011

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.

Dr. Cathy Savage-Dunn
Chair of Examining Committee

Dr. Edward Kennelly
Executive Officer

Dr. Christine Li
Dr. Peter Gergen
Dr. Timothy Short
Dr. Wilma Saffran
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

DBL-1/BMP target genes and *sma-9* function in *C. elegans*

by

Jianghua Yin

Adviser: Professor Cathy Savage-Dunn

Transforming growth factor β (TGF-β) signaling regulates a large number of biological processes including cell proliferation, lineage determination, differentiation, motility, adhesion, and death. In *Caenorhabditis elegans*, the DBL-1 pathway, a BMP-related signaling pathway, regulates body size and male tail development. A Drosophila Schnurri homolog, *sma-9*, acts as a transcription cofactor in the DBL-1 pathway (Liang et al., 2003). By comparing gene expression profiles of *sma-9* and *dbl-1* mutants to wild type animals, we found that DBL-1/BMP signaling pathway regulates a large number of target genes including collagen genes, transcription factors, genes involved in innate immunity, dauer and aging genes (Liang et al., 2007). Here, we have verified a series of potential target genes and their involvement in body size regulation. T27F2.4, a bZip transcription factor, is expressed in intestine and neurons, is negatively regulated by DBL-1/BMP signaling and inhibits growth. *col-41*, a type IV cuticle collagen gene, is expressed in hypodermis, is positively regulated by DBL-1/BMP signaling and promotes normal growth. *col-41* promoter analysis indicates that the first 500bp promoter region is essential for its basal level expression, but does not contain the *cis*-regulatory elements responsible for *sma-9* regulation. The regulation of *col-41* by *sma-9* in vivo could be a combination of direct and indirect effects. In addition to T27F2.4 and *col-41*, we have also identified other target genes responsive to DBL-1/BMP signaling, which greatly

expands the potential toolkit of reporters for further analysis of DBL-1/BMP signaling.

The *sma-9* locus *in vivo* undergoes alternative splicing, including an unusual *trans*-splicing event that could generate two non-overlapping shorter transcripts. We have demonstrated that the shorter transcripts are expressed *in vivo*. Furthermore, we find that one of the short transcripts plays a tissue-specific role in *sma-9* function, contributing to the patterning of male-specific sensory rays, but not to the regulation of body size. Based on previous results, we suggest that this transcript encodes a C-terminal SMA-9 isoform that may provide transcriptional activation activity, while full length isoforms may mediate transcriptional repression.

Acknowledgements

I own my greatly thanks to my mentor Dr. Cathy Savage-Dunn for been a wonderful advisor. Without her continuous support and guidance, I could not have got this much done. During these years, she always encourages me when I feel depressed and makes me believe I am good and I can do it. Also, I would like to thank my committee members for their generous support for reaching my goal throughout my Ph.D. In addition, I want to thank Dr. Alicia Meléndez for her valuable suggestions and help, Dr. Areti Tsiola for her help on using microscopes.

I want to thank all the members in Dr. Savage-Dunn's and Dr. Meléndez's lab for their help. In particularly, I greatly thank Ling Yu for teaching me a lot of techniques in worm studies. I thank Jun Liang, Thilini Fernando and Sheng Xiong for helpful discussions on my projects.

I want to thank the Biochemistry Program at Graduate Center and the Biology Department for providing me a great environment to accomplish my research.

Finally, I want to thank my family for unconditional support, love and encouragement. Especially thank my brother at home for taking good care of our parents. I also want to greatly thank my girlfriend for her encouragement and support outside the lab.

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

I.1 *C. elegans*: a classic genetic model organism

In 1965, Sydney Brenner first introduced *Caenorhabditis elegans* as a model system to study animal genetics, development and behavior. However, it rapidly turned to be a popular tool for scientists to study almost all fundamental biological processes, including embryonic and post-embryonic development, neurobiology, apoptosis, signal transduction, and evolution as well. Meanwhile, it was also used to build disease models to uncover underlying mechanisms, seeking for potential clinical therapy and drug discovery.

Compared to other well developed model organisms, such as Drosophila, zebrafish and mice, *C. elegans* offers a lot of advantages for genetic studies, especially for understanding developmental processes. It has a rapid life cycle, transparent body and is easy for laboratory cultivation. In addition, its relatively small genome, well defined cell lineage and easy transformation add credits to its popularity for various study purposes.

I.1.1. Rapid life cycle.

C. elegans has a rapid life cycle: around 3 days for each generation and each single hermaphrodite can give rise to 300-350 progeny. Normal *C. elegans* has six developmental stages: eggs, L1, L2, L3, L4, and adult (Fig. 1.1). At 22°C, embryogenesis requires approximately 14 hours and the post-embryonic development from first larval stage (L1) stage to adult needs 36 hours. In each developmental stage, a stage-specific cuticle is secreted by the hypodermis. As a result, molting the old cuticle and forming a

new one is the indication of stage change.

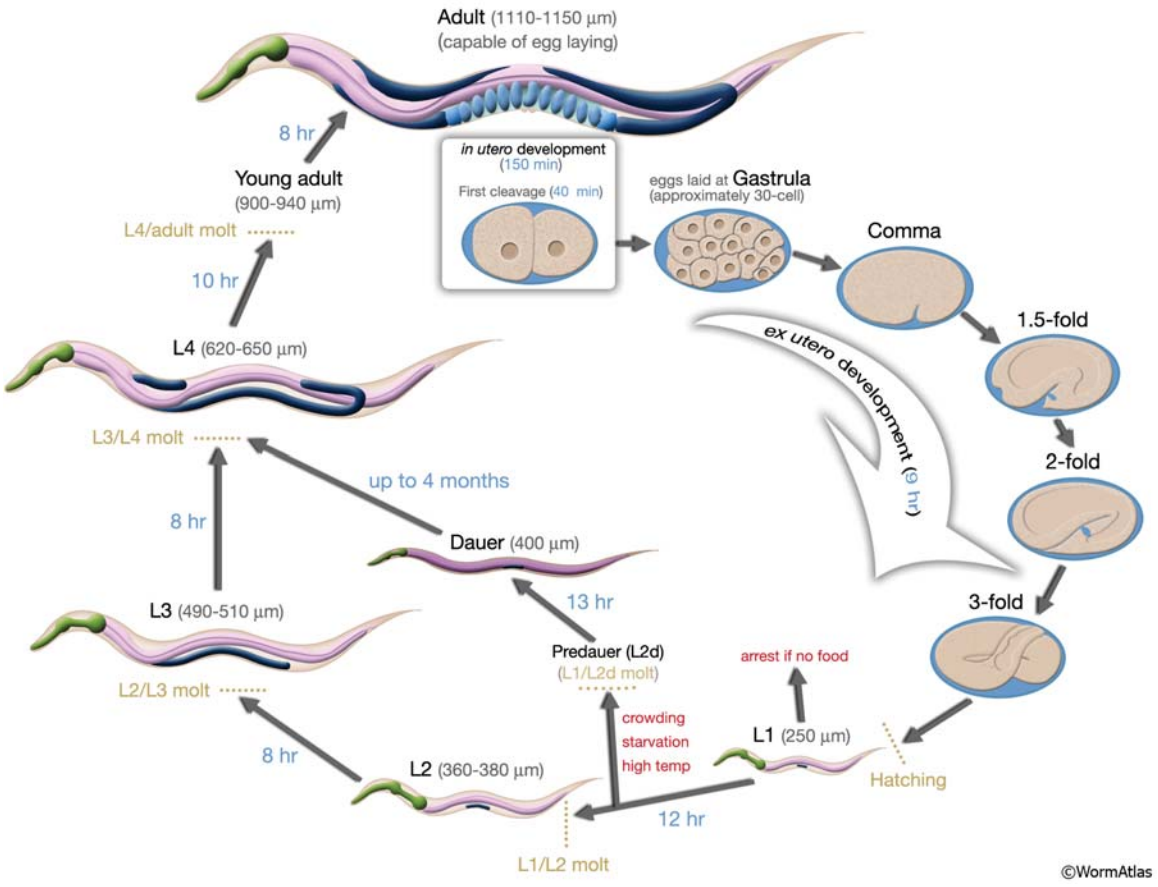


Figure 1.1 Life cycle of *C. elegans* at 22°C. (WormAtlas.org)

Under unfavorable living environments, for example, short of food, high temperature and overcrowded growth, worms in L2 stage would go into a non-developmental stage, the dauer stage, instead of the L3 stage. Worms that enter the dauer stage change their energy metabolism, accumulating fat and altering their behavior strikingly. Also, a specific cuticle replaces the one for L2. Dauer larvae can survive for months, approximately ten times their normal life span. When the environment changes back to favorable conditions, they resume development into the L4 stage, bypassing L3 stage. *C. elegans* has a transparent body throughout development. This is a great advantage for us to examine cells *in vivo* and mutant phenotypes. In addition, the cell

number and their position are consistent between individuals. Adult worms have less than 1000 cells (excluding germ cells), including 302 neurons in hermaphrodites or 381 neurons in males. The entire cell lineage, from egg to adult, is essentially the same in each animal and is known precisely (Sulston et al. 1977, 1983). This benefits studies on cell fate determination in embryonic and post-embryonic development.

The hermaphrodites and rare males are another key characteristic for *C. elegans*. As a result, self fertilization and cross fertilization are co-existent in *C. elegans*. The hermaphrodites are protandrous, first producing sperm in late L4 stage, storing them in its spermathecae, and then turning to produce oocytes. Rare males can mate with hermaphrodites, stimulating the production of oocytes by hermaphrodites. Thereby, homozygotic and heterozygotic animals can be easily obtained from mutagenesis or cross. Meanwhile, easy transformation of *C. elegans* by microinjection makes them suitable for genetic analysis, and one can knockdown specific genes using dsRNA through both microinjection and feeding. Furthermore, *C. elegans* has a small genome, about 97 million base pairs. Approximately, 19,099 genes (The *C. elegans* Sequencing Consortium 1998) are predicted and only 6% of them have been studied genetically or biochemically. There are still a large number of genes whose functions are unknown. This provides a challenging space for researchers to explore.

I.1.2. Anatomy of *C. elegans*

Basically, like other nematodes, the body of *C. elegans* is made up of two concentric tubes separated by a fluid-filled space, the pseudocoelom. The animal's shape is maintained by internal hydrostatic pressure. The outer tube contains the nervous

system, gonad, coelomcytes, and excretory/secretory system and is covered by the collagenous, extracellular cuticle, which is secreted by the underlying hypodermis. The inner tube is composed of the muscular pharynx with its nearly autonomous nervous system, intestine, rectum and anus (Fig. 1.2).

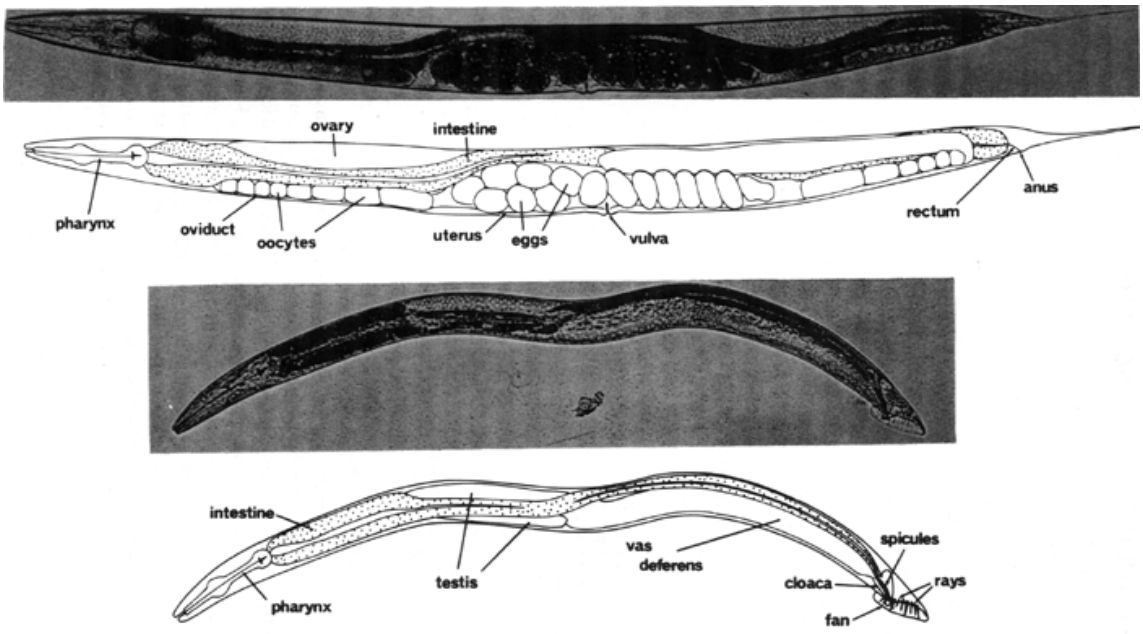


Figure 1.2 Photomicrographs showing major anatomical features of the *C. elegans* adult hermaphrodite (top) and male (bottom). Shown are lateral views under bright-field illumination. (Sulston & Horvitz. 1977).

The copulatory bursa in the male tail includes two spicules, a fan, and 18 rays. The fan is an extended acellular structure of cuticle matrix. The rays are distinguished into two identical groups that are one-axis symmetrical along the ventral midline. Each group contains nine different rays numbered 1 through 9 from anterior to posterior (Nguyen et al. 1999). Each ray contains two neurons, RNA and RNB that express different neurotransmitters, and one support cell surrounded by a hypodermal sheath. For example, dopamine is expressed only in the ray5A, ray 7A, and ray 9A neurons (Lints &

Emmons 1999). 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 Sma/Mab mutants have defects in the male tail, which will be discussed in detail later.

I.2 Transforming growth factor-β (TGF-β) signaling pathway

The transforming growth factor β (TGF-β) superfamily comprises structurally related polypeptide growth factors, including TGF-βs, bone morphogenetic proteins (BMPs), growth and differentiation factors (GDFs), Activins and Nodal (Feng and Derynck, 2005). Each member is capable of regulating a large number of biological processes including cell proliferation, lineage determination, differentiation, motility, adhesion, and death. These TGF-β superfamily ligands, as well as their downstream pathway components, are well conserved during evolution and the signal transduction pathway is relatively simple and linear. Active TGF-β ligand dimer brings together two pairs of transmembrane receptor serine/threonine kinases known as type II and type I receptors. Upon binding TGF-βs, type II receptor phosphorylates and activates type I receptor. The intracellular signal transducers (R-Smads) then are phosphorylated by type I receptor, and in turn, form complexes with Co-Smad. Those Smad complexes accumulate in the nucleus, with the help of other transcriptional cofactors, modulating transcription of target genes (Fig. 1.3). Basically, TGF-β and activins signal their transcription responses through Smad2 and Smad3, whereas BMPs signal through Smad1, Smad5 and Smad8 (Feng and Derynck. 2005).

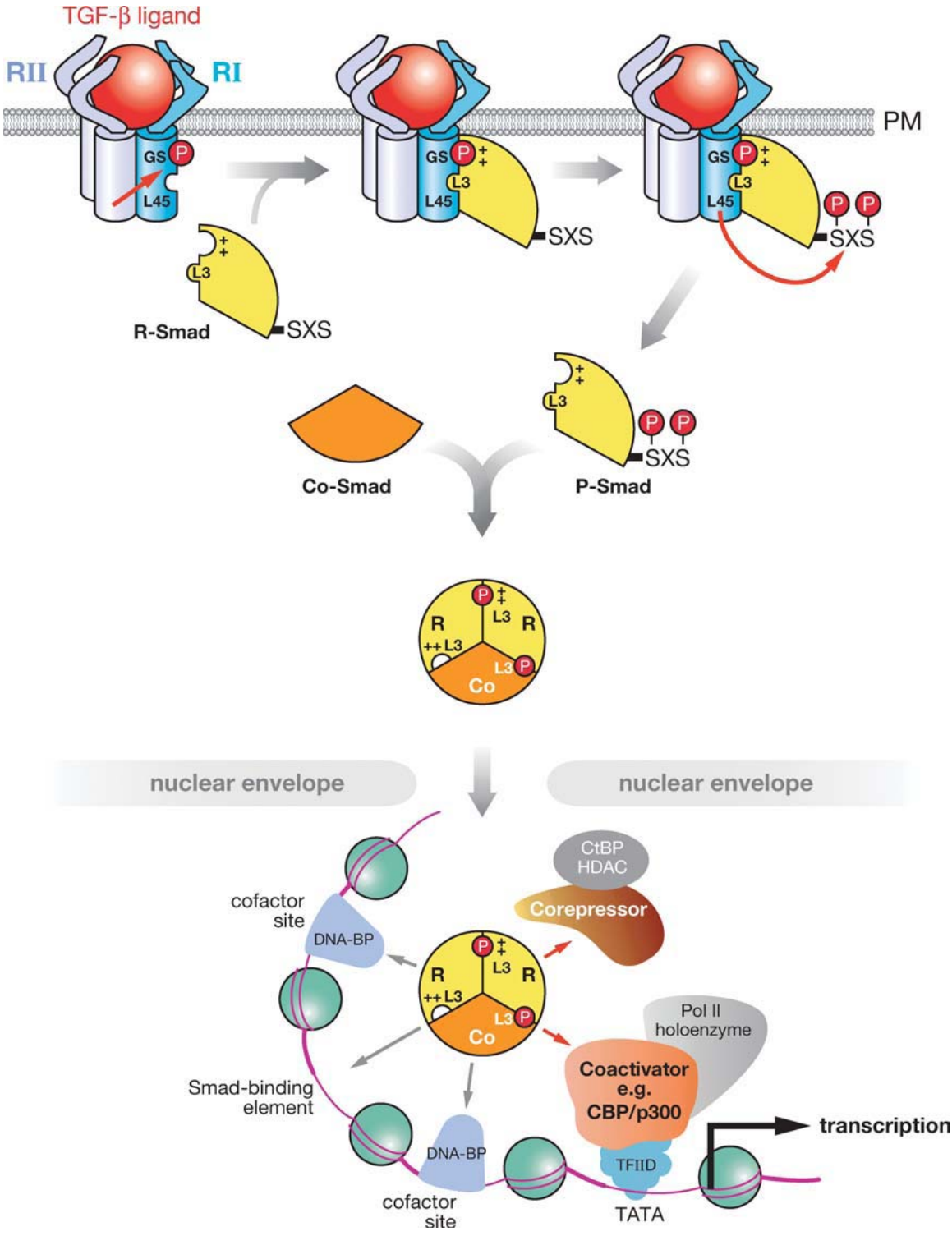


Figure 1.3 **Transforming growth factor-β (TGF-β) signaling pathway** (Feng and Derynck. 2005).

I.2.1. TGF-β ligands

Based on structural and functional relatedness, the large TGF-β superfamily divides into two subfamilies: the TGF-β isoforms, Nodal, and Activins; and the BMPs and GDFs. The TGF-β superfamily comprises more than 30 genes in mammals, 7 genes in Drosophila and 5 genes in *C. elegans*. The ligands are expressed and secreted into the extracellular matrix (ECM) as dimeric pre-proproteins. Dimerization and stabilization require the pro-domain (Gray and Mason. 1990). The mature ligands are cleaved from the prodomain by furin-like convertase (Dubois et al. 1995). Most frequently, ligand dimerizes with itself. However, heterodimerization is also observed, for example in mammals between Nodal and BMP4 or BMP7 (Yeo and Whitman. 2001), and in Drosophila, between Dpp and Scw (Shimmi et al. 2005). A major regulatory step in TGF-β signaling is the regulation of ligand accessibility by extracellular ligand binding proteins.

I.2.2. TGF-β receptors

Compared to the large number of TGF-β ligands, there are only five type II receptors (TβRII) and seven type I receptors (TβRI) identified in humans. They are transmembrane serine/threonine kinases, which contain a short extracellular region (Wells et al. 1997), a transmembrane domain and a cytoplasmic serine/threonine kinase domain. TGF-β binding stabilizes the interaction of the TβRII dimer with two TβRI molecules, forming heterotetrameric, active receptor complexes (Feng and Derynck. 2005). The assembly of the active receptor complexes is different for different ligands. In the case of TGF-β receptors, ligand-bound high affinity type II receptor will recruit type I

receptor and this process is facilitated by direct interaction between type II and type I receptor. However, the BMP receptors do not interact directly. They are linked via the ligands (Groppe et al. 2008).

Considering the relatively small number of TGF-β receptors, it is becoming apparent that different ligand-receptor interactions and type II and type I receptor pairings exist (Feng and Derynck. 2005), which can possibly account for the huge diversity of the TGF-β signaling outputs. Indeed, in endothelial cells, TGF-β can activate both ALK1 and ALK5, and a complex has been proposed comprising these two distinct type I receptors (Goumans et al. 2003). Recently, mixed receptor complexes containing TβRII, ALK5, and either ALK2 or ALK3 have been reported to mediate a novel branch of TGF-β signaling in epithelial cells (Daly et al., 2008). In Drosophila, a receptor complex comprising the type I receptor Thickveins (Tkv) and Saxophone (Sax) and the type II receptor Punt have also been observed (O'Connor et al. 2006).

Upon ligand binding and receptor complex formation, the cytoplasmic kinase domain of type II receptor will phosphorylate the GS domain on type I receptor, which in turn activates type I receptor kinase activity. Then the activated type I receptor will phosphorylate Smad proteins and transduce the signal downstream. The nine amino acid L45 loop in the kinase domain of type I receptor is the critical motif that contributes to the receptor-Smad protein interaction. Huse et al. showed from the crystal structure of type I receptor that the L45 loop is accessible for protein interactions (Huse et al. 2001). Meanwhile, receptors with different signaling specificity have different L45 sequences (Chen et al. 1998; Feng & Derynck. 1997).

I.2.3. Smad proteins

Smad family proteins are the well established intracellular effectors in TGF-β signaling. The founding member of the Smad family is the product of the *Drosophila* gene *Mad* (*mothers against dpp*) (Sekelsky et al. 1995), followed by three *Mad* homologues identified in *C. elegans*, called *sma-2*, *sma-3*, and *sma-4* (Savage et al. 1996). Later on, many homologues were discovered in vertebrates and named Smad (merge of MAD and SMA). Smads exist in three subgroups: (1) receptor-regulated Smads (R-Smads), e.g. Smad1, 2, 3, 5, and 8 in vertebrates, Mad in *Drosophila*, SMA-2 and -3, and DAF-8 and -14 in *C. elegans*; (2), common Smads (Co-Smads) e.g. Smad4 in vertebrates, SMA-4, and DAF-3 in *C. elegans*; (3) inhibitory Smads (I-Smads), e.g. Smad6 and 7 in vertebrates and Dad in *Drosophila*.

R-Smads and Co-Smads contains highly conserved MH1 (N-terminal) and MH2 (C-terminal) domains linked by a more diverse linker region. The MH1 domain has approximately 130 amino acids and contributes to the DNA-binding activity. R-Smads can directly interact with type I receptor through the MH2 domain and have a C-terminal SXS motif in which two serines are direct targets for phosphorylation by type I receptor. Phosphorylation of SXS motif in R-Smads leads to a conformational change, resulting in dissociation of R-Smad from type I receptor and formation of a trimeric Smad complex consisting of two R-Smads and a Co-Smad. This trimeric Smad complex then accumulates in the nucleus and regulates gene expression, positively or negatively, usually in association with other transcriptional cofactors (Moustakas & Heldin. 2009; Feng & Derynck. 2005; Derynck & Zhang. 2003; Shi & Massague. 2003).

Since R-Smads localize in the cytoplasm, efficient R-Smad recruitment and activation in response to TGF-β and activin requires other accessory proteins. SARA (Smad Anchor for Receptor Activation), an FYVE domain-containing, plasma membrane localized and EEA-1 positive endosome enriched protein can interact with type I receptor and Smad2/3 and recruits Smad2 to the receptor for phosphorylation (Wu et al. 2000). Thus, complex formation of the receptors with Smads and SARA at EEA-1 positive endosome can promote TGF-β signaling (Di Guglielmo et al. 2003). Hgs, another FYVE domain protein involved in endosomal trafficking, may play a role similar to that of SARA, since it also interacts with Smad2 and Smad3 and enhances ligand induced Smad phosphorylation and gene expression (Miura et al. 2000).

Inhibitory Smads (I-Smads), like Smad6 and Smad7 which are induced by TGF-β family ligands, inhibit TGF-β signaling through various mechanisms (Hayashi et al. 1997; Imamura et al. 1997; Nakao et al. 1997). For example, Smad7 is shown to form a stable complex with type I receptors, therefore leading to inhibition of R-Smad phosphorylation and the hetero-complex formation between R-Smads and Co-Smad (Hayashi et al. 1997). Smad7 can also recruit the HECT type of E3 ubiquitin ligases, Smurf1 and Smurf2 to the activated type I receptor ALK5/TbRI, leading to the degradation of the receptor through the proteasomal pathway (Ebisawa et al. 2001; Kavsak et al. 2000).

I.2.4. Smad DNA-binding and Interaction with DNA-binding transcription factors

Smad complexes can bind specific DNA sequences in the promoters or enhancers of target genes. The Smad binding element (SBE) was first identified in the PAI1 promoter, in which Smad3-Smad4 complex contacts DNA at 5'-GTCTAGAC-3' as the

optimal sequence (Zawel et al. 1998). Later, a single Smad binding site, which is contacted by the MH1 domain of Smad3 or Smad4, is defined as the sequence 5′-AGAC-3′ or its reverse complement 5′-GTCT-3′ (Shi et al. 1998). The SBEs often contain direct or inverted repeats of this core sequence.

In contrast, Smad1 and its Drosophila homolog Mad bind to a GC-rich sequence with higher affinity. The consensus for Mad binding is 5'-GNCGNC-3' (Gao et al. 2005), which tends to occur adjacent to an AGAC or GTCT element, allowing Medea (Smad4 homolog in Drosophila) to make additional contact with DNA. The spacing between these two sites is normally 5 base pairs. The length of spacing is critical for the recruitment of Schnurri (a transcriptional regulator) to the DNA-binding Mad-Medea (Mad/Med) complex (Pyrowolakis et al. 2004; Yao et al. 2006). In Drosophila, Schnurri (Shn) represses expression of the *brinker* gene via the formation of a Shn/Mad/Med complex on defined elements in the *brinker* promoter (Pyrowolakis et al. 2004). However, in mammalian cells, the Shn/Smad complex acts as an activator. These results indicate that the Shn/Smad complex may recruit co-activators or co-repressors in a context-dependent manner.

Since the short SBEs only allow low-affinity binding, other DNA-binding factors are frequently required for the specific recruitment of Smads to their specific response elements on target gene promoters. Numerous transcription factors have been shown to interact with different Smad proteins (reviewed in Feng & Derynck. 2005 and Moustakas and Heldin. 2009).

FoxH1, a forkhead (Fox) transcription factor and also a principal DNA-binding component in activin-responsive factor (ARF), is the first transcription factor reported to

interact and cooperate with Smads in mediating TGF-β signaling (Chen et al. 1996, 1997). In response to activin, Smad2/4 complexes interact with DNA-bound FoxH1 at an activin-response element and provide ligand-induced transcription (Chen et al. 1997). In this complex, Smad4 contacts the DNA while the MH2 domain of Smad2 interacts with FoxH1 at the common Smad interaction motif (SIM) and FoxH1 motif (FM) (Labbe et al. 1998; Zhou et al. 1998; Randall et al. 2002). Interestingly, another group of Smad2-interacting transcription factors, the Mix family of homeodomain proteins, shares the same Smad2-interaction motif with FoxH1, but contain a distinct DNA-binding domain and exhibit a different expression pattern. Therefore, they induce a different set of nodal-responsive genes compared to that regulated by FoxH1 (Germain et al. 2000; Kunwar et al. 2003).

Another forkhead transcription factor, FoxO, can interact with Smad3/4 complex to induce transcription of the CDK inhibitor *p21CIP1* gene (Seoane et al. 2004). Gomis et al. further identify a synexpression group that is transcriptionally induced by TGF-β through the FoxO–Smad combination. This FoxO–Smad synexpression group contains 11 genes, including *p21CIP1,* which define the cytostatic, apoptotic and adaptive signaling responses of keratinocytes to TGF-β.

FoxO binds to a distal sequence of the *p21CIP1* gene promoter, however, Smads can also interact with Sp1 (a zinc finger transcription factor) at a proximal sequence to induce *p21CIP1* gene expression (Pardali et al. 2000), which indicates the possibility of multiple Smad complex formations in a single gene promoter with the involvement of different transcription factors. Besides the *p21CIP1* gene, Smad-Sp1 interaction can also induce other TGF-β target genes, including *p15INK4B*, Smad7, PAI-1, and collagen

(Feng et al. 2000; Pardali et al. 2000). Activation of *Vent2* during ventral mesoderm differentiation in Xenopus is another example of zinc finger transcription factor participating in BMP signaling. OAZ, which contains 30 zinc finger motifs, can associate with Smad1/4 complex and bind to *Vent2* promoter sequence, inducing *Vent2* expression in response to BMP (Hata et al. 2000).

Besides OAZ, another DNA-binding partner for BMP-regulated R-Smads is Drosophila Schnurri. Shn contains seven C_2H_2 zinc finger motifs organized into three clusters, localizes in the nucleus and is able to form a complex with the *Drosophila* BMP-regulated R-Smad and Co-Smad, Mad and Medea on target gene promoters to activate or repress transcription in response to Dpp. Brinker, a transcriptional repressor, is an important target gene that is repressed by the Shn/Mad/Med complex, which in turn, derepresses a large number of *dpp* target genes (Marty et al. 2000). The transcriptional repression of *brinker* solely relies on the capacity of a Dpp-dependent silencer element (SE) to interact with Mad and Medea and to subsequently recruit Shn as a scaffolding protein. This Shn/Mad/Med -dependent silencer also functions to repress some other key genes in development, such as *bam* and *gsb* gene (Pyrowolakis et al. 2004; Yao et al. 2006). Additionally, whether Smad/Shn complex behaves as a transcriptional repressor or activator does not depend on the sequence of the cis-regulatory element within the promoter. Smad/Shn complexes are likely to recruit coactivators and corepressors in a context-dependent manner (Yao et al. 2006). Furthermore, it has been demonstrated in Drosophila that Schnurri can bind to the enhancer of Ultrabithorax (Ubx) and activate its expression in response to Dpp signaling (Dai et al. 2000).

Smads also interact with select bZIP family transcription factors, which contain

basic and leucine zipper domains involved in DNA binding and dimerization. One example is ATF3, which can interact with Smad3 to inhibit Id (inhibitor of differentiation) genes in response to TGF-β (Kang et al. 2003). Analysis of the Id1 promoter has identified the ATF-binding site as an element specifically required for TGF-β-induced repression. In fact, ATF3 expression is induced in response to TGF-β, but not BMP and can not interact with Smad1 either. Thus, a Smad3-mediated primary gene response, ATF3 induction, enables Smad3 to participate in an ATF3-mediated, secondary gene response (Kang et al. 2003). This "self-enabling" TGF-β response describes the scenario that the transcription factor interacting with Smad protein is actually synthesized in response to the signal. Similar examples are also found in other organisms. In Xenopus, Nodal signaling can induce expression of the transcription factor Mixer and then Mixer will cooperate with activated Smads to regulate *Goosecoid* expression (Germain et al. 2000). In the Drosophila embryo, peak levels of Dpp signaling induces *zerknüllt (zen)* expression in the dorsalmost region, which then combines with Smads, binds to the enhancer of *Race* (related to angiotensin-converting enzyme) and activates its expression (Xu et al. 2005). Other members of the bZIP family are also found to interact with Smads. Smad3 can interact with c-Jun, JunB, ATF-2, and with lower efficiency, c-Fos (Liberati et al. 1999; Sano et al. 1999; Zhang et al. 1998).

I.2.5. Coactivators and corepressors

In addition to interaction with DNA-binding transcription cofactors, Smad complexes often recruit coactivators or corepressors to fulfill TGF-β mediated transcriptional activity.

CBP and p300 are two early-known coactivators recruited by R-Smads. They function as Smad4-dependent transcriptional coactivators for Smad3 and associate with R-Smads through the MH2 domain (Feng et al. 1998). CBP and p300 could recruit the basal transcriptional machinery including RNA polymerase II to the promoter, and could also use their intrinsic histone acetyltransferase (HAT) activity to modify chromatin structure. Furthermore, the linker region of Smad3 contains a transcriptional activation domain and its activity requires the binding of p300 in the linker domain (Wang et al. 2005). Activator-recruited co-factor (ARC) complex or the metazoan Mediator complex acts as a coactivator in transcription through its interaction with RNA polymerase II (Naar et al. 2001; Rachez & Freedman. 2001). ARC105, a component of this complex, is recruited to the Smad-responsive promoter *mix.2* and *gsc* gene in response to activin/nodal and binds Smad2/3 and Smad4, but not Smad1, in response to TGF-β (Kato et al. 2002). Thus, the Smad-ARC105 interaction mediates and relays TGF-β signaling to the Pol II machinery, which activates select genes.

Corepressors that directly bind to Smads repress transcription induced by Smads. Ski/SnoN, nuclear oncoproteins, can interact directly with Smad2, Smad3, and Smad4 on a TGF-β-responsive promoter element and repress their abilities to activate transcription (Luo et al. 1999). Binding of Ski/SnoN to Smad4 interferes with the interaction between Smad4 and phosphorylated R-Smads, disrupting the active heteromeric complex formed between the Co-Smads and the R-Smads (Luo et al. 2004). Another oncoprotein, c-Myc, associates with Smad2/3 and does not interfere the formation of a Smad-Sp1 complex. Binding of c-Myc to the Smad-Sp1 complex inhibits TGF-β-induced transcriptional activity of Sp1 and Smad/Sp1-dependent transcription of the p15Ink4B gene (Feng et al.

2002). TGIF is a homeobox transcription factor and can interact with Smads to repress Smad-mediated transcription. It recruits histone deacetylases (HDAC) and competes with CBP/p300 for R-Smad interaction (Wotton & Massague. 2001; Wotton et al. 2001).

I.3 TGF-β signaling pathway in *C. elegans*

In *C. elegans*, five TGF-β-related genes can be identified by sequence homology: *daf-7*, *dbl-1*, *unc-129*, *tig-2*, and Y46E12BL.1. No biological roles have yet been described for *tig-2* or Y46E12BL.1. *daf-7* is a TGF-β-related ligand and *dbl-1* is more related to BMP2 and Dpp. Their downstream signaling components have also been characterized, referred to as the Dauer pathway and the Sma/Mab pathway (Fig. 1.4), respectively (Suzuki et al. 1999; Savage-Dunn. 2005).

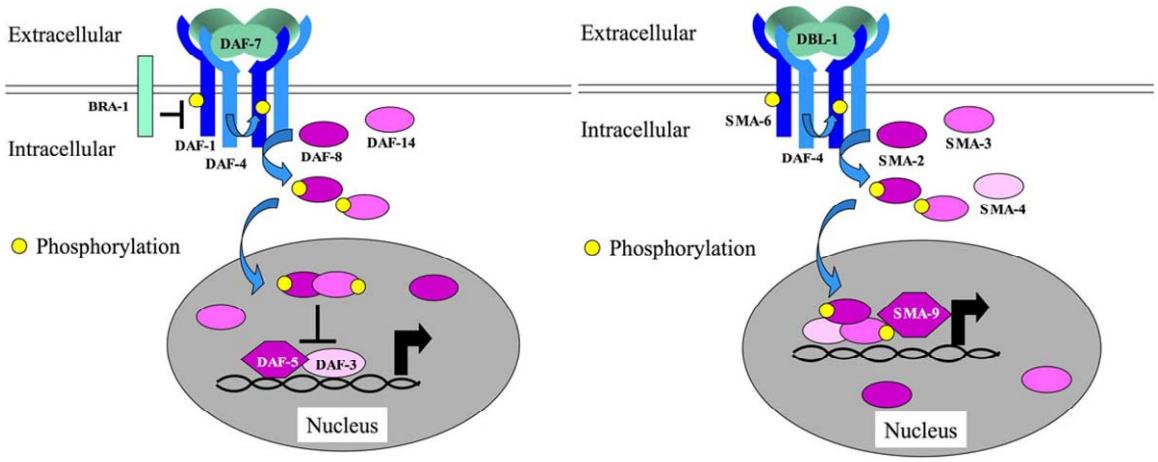


Figure 1.4 Dauer pathway and Sma/Mab pathway in *C. elegans* (Savage-Dunn. 2005).

I.3.1. The Dauer pathway

Under unfavorable living conditions (high population density, low food

availability, high temperature), the worms will go into the dauer diapause as mentioned before. The dauer TGF-β pathway is one signaling cascade that regulates this process. DAF-7 acts as the ligand (Ren et al. 1996), which binds to type II receptor DAF-4 (Estevez et al. 1993) and type I receptor DAF-1 (Georgi et al. 1990). Upon ligand binding, DAF-1 phosphorylates R-Smad DAF-8 and DAF-14 (Riddle & Albert. 1997; Inoue & Thomas 2000). Since DAF-8 and DAF-14 do not have identifiable DNA binding domains, they are unlikely to transduce signals directly into the nucleus. In contrast, they negatively regulate Co-Smad DAF-3 (Patterson et al. 1997), which associates with the Ski/SnoN homolog DAF-5 (da Graca et al. 2003) to form a transcription factor complex to control dauer entry (Liu et al. 2004).

I.3.2. Sma/Mab pathway

The Sma/Mab pathway was first characterized by Savage et al. in 1996 and regulates body size and the development of male-specific sensory rays and copulatory spicules (Savage et al. 1996). Mutants of components in this pathway have small body size (Sma) phenotype and abnormal male tail (Mab) phenotype. The first three components in this pathway, *sma-2*, *sma-3* and *sma-4* were cloned by this group (Savage et al. 1996). Shortly after, other related components are discovered. For instance, DBL-1 acts as the ligand (Suzuki et al. 1999) and SMA-6 acts as the type I receptor (Krishna et al. 1999). They share the same type II receptor DAF-4 with the Dauer pathway.

Mutations in Sma/Mab pathway components cause the Sma phenotype. *dbl-1* pathway mutants show decreased seam cell length but normal seam cell nuclei number,

suggesting that decreases in cell size rather than cell number are responsible for *dbl-1* mutant phenotype (Wang et al. 2002). Organ size measurements have shown that different organs are reduced in size to different degrees. The degree of reduction in the size of hypodermal seam cells and hyp7 is most proportional to the degree of reduction in body size (Wang et al. 2002; Nagamatsu and Ohshima. 2004), consistent with the defined focus of action of Sma/Mab signaling components in the hypodermis. Loss of *dbl-1* activity results in smaller animals, while *dbl-1* overexpression results in longer animals, suggesting that *dbl-1* acts as a dose-dependent regulator of body size (Morita et al. 1999). In addition, comparison of the growth curves between the wild type and mutant animals shows that the mutant animals have the same body size as wild type at the L1 stage. However, the adult animals show only around 50% body size of the wild type (Savage-Dunn et al. 2000), suggesting that a post-embryonic defect causes the Sma phenotype.

In addition to the Sma phenotype, a Mab phenotype is also found in these mutants. Male tail sensory rays and spicules are essential for mating. Normally in wild type animals, there are nine rays on each side of the male tail. However, *dbl-1*, *daf-4*, *sma-2*, *sma-3*, *sma-4*, and *sma-6* mutant males show abnormal ray fusions and crumpled spicules. Ray fusion frequently occurs between rays 4-5, 6-7, and 8-9 (Savage et al. 1996).

In addition to body size regulation and male tail development, the Sma/Mab pathway has been shown to regulate innate immunity of *C. elegans*. *sma-2*, *-3*, *-4* and *-6* are required for resistance to *Pseudomonas aeruginosa.* The loss of function mutants of those four genes are hypersensitive to infection (Mallo et al. 2002). Liu's lab shows mutations in the Schnurri homolog *sma-9* cause ventralization of the M lineage and that wild-type SMA-9 antagonizes the Sma/Mab TGFbeta pathway to promote dorsal M

lineage fates (Foehr et al. 2006). More recently, Murphy's lab shows the involvement of Sma/Mab signaling in reproductive aging. Reduction of Sma/Mab signaling delays reproductive aging by maintaining oocyte and germline quality. Compared to cell-autonomous regulation of body size development, Sma/Mab signaling regulates oocyte and distal germline quality maintenance nonautonomously and this process is temporally and transcriptionally separable from its regulation of growth (Luo et al. 2009; Luo et al. 2010).

To date, few other proteins are found to interact with Sma/Mab pathway components. Padgett's lab showed that LON-2 can negatively regulate Sma/Mab pathway activity. LON-2 is a conserved member of the glypican family of heparan sulfate proteoglycans and can bind BMP2 in vitro. *lon-2(lf)* mutations result in a long body size phenotype and the Drosophila glypican gene *dally* rescues the *lon-2(lf)* body size defect (Gumienny et al. 2007). More recently, the sole member of the repulsive guidance molecule (RGM) family of proteins in *C. elegans*, DRAG-1, is found to positively regulate Sma/Mab pathway activity at the ligand-receptor level (Tian et al. 2010).

I.3.3. *sma-9* in TGF-β signaling pathway in *C. elegans*.

sma-9 was isolated from forward genetic screens for additional mutations affecting body size (Savage-Dunn et al. 2003). *sma-9* mutant animals show both Sma and Mab phenotypes, the same as other Sma/Mab pathway component mutants. The sequence analysis shows that the predicted *sma-9* gene encodes a Zinc finger transcription factor

homologous to Shn in Drosophila. As mentioned before, Shn act as a Mad cofactor in the Dpp pathway, and is essential for Dpp-mediated repression of brinker gene transcription (Marty et al. 2000). SMA-9 contains seven C_2H_2 zinc finger motifs at the C terminus clustered into two pairs and one triplet. Two acidic residue-rich domains (ARD) may correspond to transcriptional activation domains. The presence of four nuclear localization sequences (NLSs) suggests that *sma-9* functions in the nucleus. This is confirmed by Anti-SMA-9 antibody staining (Liang et al. 2003). Based on these data, we consider that SMA-9 functions as a Smad transcription cofactor to confer specific responses to DBL-1/BMP-related pathway signaling to regulate body size and male tail development.

Our recent work shows that SMA-9 functions mainly as a transcriptional repressor in body size regulation. From *sma-9*-regulated genes identified by microarray, there are more repressed genes than activated genes (Liang et al. 2007). However, emerging evidence causes us to consider that SMA-9 may not simply act as a repressor. It is possible that different domains of SMA-9 may have different transcriptional activities, whether activator or repressor. In my thesis work, I try to address what the DBL-1/BMP target genes are and how they regulated by the DBL-1 signaling, and how they are involved in body size regulation in *C. elegans*. Meanwhile, I also addressed how different *sma-9* splicing variants function in body size development and male tail patterning.

Chapter II. DBL-1/BMP target genes in *C. elegans*.

(Contributions: JY and CSD designed, performed most of the experiments and interpreted the data. Edlira Yzeiraj performed RNAi and real-time RT-PCR on collagen genes.)

II.1 Abstract

In *Caenorhabditis elegans*, the DBL-1 pathway, a BMP-related signaling pathway, regulates body size and male tail development. A Drosophila Schnurri homolog, *sma-9*, acts as a transcription cofactor in the DBL-1 pathway (Liang et al. 2003). By comparing gene expression profiles of *sma-9* and *dbl-1* mutants to wild type animals, we found that DBL-1/BMP signaling regulates a large number of target genes including collagen genes, transcription factors, genes involved in innate immunity, dauer and aging genes (Liang et al. 2007). Here, we have verified a series of potential target genes and their involvement in body size regulation. T27F2.4, a bZip transcription factor, is expressed in intestine and neurons, is negatively regulated by DBL-1/BMP signaling and inhibits growth. *col-41*, a type IV cuticle collagen gene, is expressed in hypodermis, is positively regulated by DBL-1/BMP signaling and promots normal growth. *col-41* promoter analysis indicates that the proximal 500bp upstream promoter region is essential for its basal level expression, but does not contain the *cis*-regulatory elements responsible for *sma-9* regulation. The regulation of *col-41* by *sma-9* in vivo could be a combination of direct and indirect effects. In addition to T27F2.4 and *col-41*, we have also identified other target genes responsive to DBL-1/BMP signaling, which greatly expands the potential toolkit of reporters for further analysis of DBL-1/BMP signaling.

II.2 Introduction

The bone morphogenetic proteins (BMPs) are members of the transforming growth factor β (TGF-β) superfamily of secreted peptide growth factors that regulate a critical array of cellular processes including cell proliferation, lineage determination, differentiation, motility, adhesion, and death (Massague J. 1998). Signal transduction for this family is initiated by formation of the active ligand-receptor complex between TGF-β ligand and two transmembrane ser/thr kinase receptors, and further transduced through the intracellular Smad proteins. Smads function by forming Smad complexes consisting of two phosphorylated R-Smads and one Co-Smad that accumulate in the nucleus and regulate downstream target gene expression. Considering the relatively small number of receptors and Smads in this canonical signaling cassette, how TGF-β signaling establishes the specificity of target gene regulation to fulfill a wide array of biological responses is not well understood. In vitro, Smad complexes can bind DNA, however, to a low complexity sequence with low affinity, suggesting that, in vivo, efficient and specific promoter binding and target gene regulation require interaction with additional transcription cofactors.

In the nematode *C. elegans*, a BMP-related signaling pathway regulates body size and patterning of sex-specific tissues of male posterior (Patterson and Padgett. 2000; Savage-Dunn. 2001). This pathway, which we will refer to as the DBL-1/BMP pathway, consists of 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

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.

The *sma-9* gene, encoding a transcription cofactor of the DBL-1 pathway which is required for body size and male tail patterning, was previously identified by our lab using a genetic approach (Liang et al. 2003). *sma-9* is also required for patterning of the mesodermal lineage, in which it acts antagonistically to the DBL-1 pathway (Foehr et al. 2006). *sma-9* is predicted to encode a zinc finger transcription factor homologous to *Drosophila* Schnurri. Schnurri (Shn) functions in the BMP related *dpp* pathway in Drosophila and has been extensively studied. It is recruited to the Silencer Element (SE) of the *brinker* gene promoter by Mad/Med complex, repressing brinker expression, which in turn, derepresses a set of target genes repressed by Brinker. Some other Dpp target genes, such as *bam* and *gsb*, are also found to be directly repressed via Shn/Mad/Med - dependent silencer (Pyrowolakis et al. 2004; Yao et al. 2006). The similar interaction between SE and Shn/Smad complex is also found in vertebrates, although with reversed outcome, suggesting the conserved regulatory mechanism of Shn family proteins in signaling pathways.

In order to investigate how the DBL-1/BMP pathway regulates its target genes in *C. elegans*, our lab previously conducted microarray analysis comparing gene expression profiles of *dbl-1* and *sma-9* mutants to wild type animals (Liang et al. 2007). The microarray results show that the DBL-1/BMP signaling pathway regulates a large number of target genes including collagen genes, transcription factors, genes involved in innate immunity, dauer and aging genes. In addition, more repressed genes are found than

activated genes, which is consistent with the importance of repression in mediating body size regulation by the *dbl-1* pathway (Liang et al. 2007). In this chapter, we validate a series of potential target genes of the *dbl-1* pathway, and investigate the mechanism of target gene selection and function primarily in body size regulation. We show that collagen genes indeed are an important target gene group regulating body size development in *C. elegans*. Among them, the *col-41* gene is transcriptionally activated by the *dbl-1* pathway and a small body size phenotype results when *col-41* is depleted using RNAi, which mimics the *sma-9* mutant body size phenotype. The fact that SMA-9 fusions to known transcriptional activators and repressors can each strongly activate *col-41* expression in *sma-9* mutants suggests that the control of *col-41* expression by *sma-9* could be the combined effect of direct and indirect regulation. T27F2.4, a bZip transcription factor, is also involved in body size regulation and is repressed by *sma-9*. T27F2.4 may be involved in relaying the activity of the DBL-1 pathway to indirectly regulated target genes in the intestine.

II.3 Materials and methods

Strains

C. elegans strains were cultured using standard methods and grown at 20℃ unless otherwise noted (Brenner, 1974). In addition to strains generated in this work, the following strains were used: N2 (wild type), LG IV: *dbl-1(wk70)*; LG V: *him-5(e1490)*, LG X: *sma-9(wk55, qc3, cc604)* (Savage-Dunn et al. 2003), *lon-2(e678), dbl-1(ctIs40)*; BC15777 (*dpy-5(e61)*; sEx15777[*fat-7*p::*gfp* + *dpy-5(+)*]); BC10350 (*dpy-5(e61); sEx10350*[*clec-52*p::*gfp* + *dpy-5(+)*]); CF1553 (*muIs84*[*sod-3*p::*gfp*]); CS332 and CS339 are described in the study of Liang et al. 2007.

Molecular cloning and sequencing

GFP reporter constructs were generated by insertion of target gene promoter regions into GFP reporter pPD117.01. The T27F2.4 transcriptional GFP reporter was generated by fusion PCR (Hobert. 2002) using 4kb T27F2.4 upstream promoter region, GFP coding region from pPD117.01 and 1.0kb genomic sequence after stop codon as the templates. A T27F2.4 translational reporter was also generated by fusion PCR using 4.0kb upstream promoter region plus T27F2.4 coding region, GFP coding region from pPD117.01 and 1.0kb genomic sequence after stop codon as the templates. The *col-41* transcriptional GFP reporter was generated by fusion PCR using 1.6kb upsteam promoter region and GFP coding region from pPD117.01 as the template. 2xNLS::mCherry transcriptional reporters for other genes were generated by inserting promoter regions of genes of interest into 2xNLS-mCherry(optimized) vector (converted from pPD95.75).

The *col-41* deletion series constructs were generated by using Phusion Site-Directed Mutagenesis Kit (NEB) and confirmed by sequencing.

All primer sequences are available upon request.

Transgenic animals

Transgenic nematodes were generated by microinjection of plasmid DNA or PCR product into the gonadal syncytia of N2 hermaphrodites, with *rol-6* (100ng/ul), *myo-2::GFP* (5ng/ul) or *myo-3::mCherry* (20ng/ul) as a marker (Mello et al. 1991). Arrays are integrated into chromosomes using ɣ-irradiation.

RT-PCR

Real time RT-PCR was performed on a LightCycler 2.0 (Roche). Data analysis software used was LightCycler software version 4.0. Animals were collected at desired time points and total RNA extracted by Trizol (Invitrogen) as described (Liang et al. 2003). SuperScriptTM III platinum two-step qRT-PCR kit with SYBR Green (Invitrogen) was used to perform RT-PCR. Actin gene *act-1* was used as standard control. A primer in an intron of the *act-1* gene was used to confirm the absence of genomic DNA in RNA preps. Primer sequences for all RT-PCR experiments are available upon request.

RNAi feeding

RNAi feeding was performed as described in (Kamath and Ahringer. 2003). 12 L4 animals were transferred to feeding plates, incubated overnight, transferred to fresh feeding plates to lay eggs for 4 hours and the stage-synchronized progeny were scored.

For the growth curve, RNAi hypersensitive *rrf-3* mutants were fed T27F2.4, *col-41*, *col-141*, or *rol-6* dsRNA.

II.4 Results

II.4.1 T27F2.4 gene, a bZip transcription factor, is regulated by the DBL/BMP signaling pathway and inhibits growth

Among the candidate target genes we identified, T27F2.4 was notable for several reasons. First, its expression is significantly upregulated in *dbl-1* and *sma-9* mutant backgrounds by 2.89 and 4.33 fold, respectively. Second, quantitative real-time PCR confirmed this expression change. Third, T27F2.4 is required for body size regulation. We previously determined the physiological function of T27F2.4 gene by microinjection of T27F2.4 dsRNA, resulting in a long body size phenotype (Liang et al. 2007). To confirm this result, I knocked down T27F2.4 gene function *in vivo* by RNAi feeding. As shown in Fig. 2.1, when T27F2.4 gene expression is depleted, a long body size phenotype is observed in *rrf-3* (RNAi hypersensitive) mutant background from the second larval (L2) stage. We next want to determine where and when the T27F2.4 gene is expressed and how it is regulated by the DBL-1/BMP pathway.

In order to examine the *in vivo* expression pattern of the T27F2.4 gene, I constructed transgenic animals carrying transcriptional or translational GFP reporters. In the wild-type background, both transcriptional and translational GFP reporter expression levels are low. They express in the intestine, as well as some neurons in the head and tail region (Fig. 2.2). In the *sma-9* mutant background, however, the T27F2.4 transcriptional reporter shows significant upregulation in the intestine, but not in the neurons (Fig. 2.2).

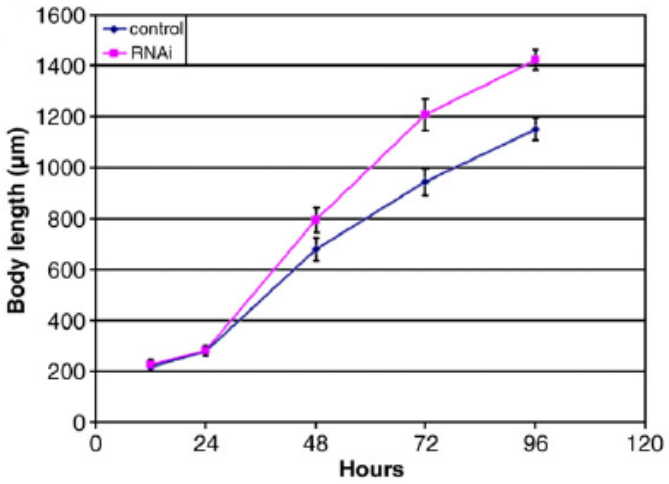


Figure 2.1 T27F2.4 inactivation by RNA interference. *rrf-3* RNAi hypersensitive animals are fed control bacteria or T27F2.4 dsRNA-expressing bacteria. RNAi animals show a long body size phenotype starting at 24 hours after egg collection (L2 stage).

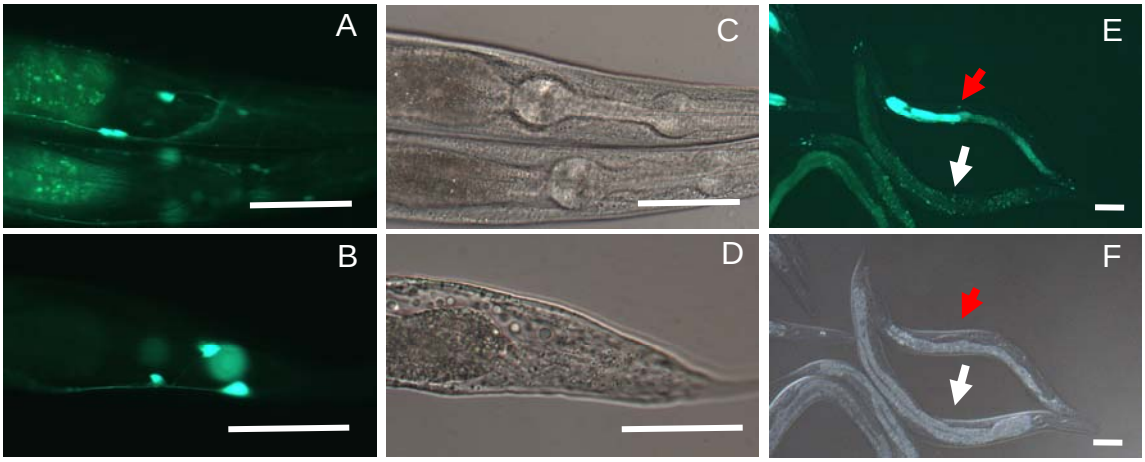


Figure 2.2 T27F2.4 transcriptional reporter expression *in vivo*. A and B, the neuronal expression of T27F2.4 in the head and tail regions of adult animals. C and D, the bright field image of the transgenic worms in A and B, respectively. E and F, T27F2.4 expression is upregulated in *sma-9* mutants. In wild-type (E, white arrow), neuronal and intestinal expression were observed in adulthood. In *sma-9(wk55)* (E, red arrow), an increased level of intestinal expression was observed in adults. (F) Bright field image of animals in panel E. Scale bars: 50um.

Interestingly, the translational reporter does not show this upregulation pattern, possibly due to post-transcriptional regulation of T27F2.4 protein levels.

The expression profile of T27F2.4 and its physiological function demonstrate that it is a valid DBL-1 target gene candidate, which regulates body size development in *C. elegans.* In wild type, DBL-1 ligand triggers the signaling through SMA-9, repressing T27F2.4 gene expression to promote normal growth. In the *sma-9* mutant background, T27F2.4 expression is derepressed, and in turn inhibits growth, resulting in a small body size phenotype.

II.4.2 COL-41, a type IV and VIII cuticle collagen, is a target gene of the DBL/BMP signaling pathway that promotes growth.

Collagens are major components in the extracellular matrix. They also have crucial functions in cell adhesion and migration. From our microarray analysis, several cuticle collagen genes show expression level changes in *dbl-1* and *sma-9* mutant backgrounds. *col-37* and *col-141* are activated in *dbl-1* mutants. In contrast, *rol-6* and *col-41* are repressed (Liang et al. 2007). We also performed real-time RT-PCR analysis to confirm the regulation of expression of these genes. As shown in Table 2.1, in both *dbl-1* and *sma-9* mutant backgrounds, *col-41* and *rol-6* expression are repressed, which is consistent with the microarray data. *col-141*, however, shows the expected increased expression in *dbl-1* but not in *sma-9* mutant background. Thus, the RT-PCR results confirm that these three collagen genes are transcriptionally regulated by DBL-1.

Table 2.1 Cuticle collagen genes expression. Real-time RT-PCR analysis of *rol-6*, *col-41* and *col-141* gene expression in wild-type, *dbl-1* and *sma-9* mutant background. Numbers under N2, *dbl-1* and *sma-9* indicate relative percentage of expression compared to control *act-1*.

	N2	*dbl-1*	Fold Change	*sma-9*	Fold Change
rol-6	**9.1±2.87**	**7.4±1.3**	**-1.23**	**3.1±0.01**	**-2.93**
col-41	**10.8±0.73**	**8.7±3.2**	**-1.24**	**5.3±0.04**	**-2.04**
col-141	**0.7±0.093**	**1.2±0.02**	**1.71**	**0.2±0.01**	**-3.50**

The involvement of cuticle collagen genes in body size has been previously demonstrated, since collagen mutants have phenotypes such as dumpy, squat, and long (Johnstone. 2000; Kramer et al. 1988, Gumienny et al. 2007; McMahon et al. 2003). Therefore, we want to know if the cuticle collagen genes from our microarray analysis are also involved in body size regulation. We tested the physiological functions of *col-41*, *rol-6* and *col-141* by RNAi feeding. As shown in Fig. 2.3, *col-41* and *rol-6* RNAi animals show significant growth defects compared to control animals, which mimics *sma-9* mutant phenotype and is consistent with their expression profile in *sma-9* mutant background. In contrast, *col-141* RNAi results in a long body size phenotype until 72 hours of development (L4 stage), and interestingly reverts to wild-type length at 96 hours (adulthood). The long body size phenotype is consistent with the transcriptional repression of *col-141* by DBL-1, and demonstrates that *col-141* is a negative regulator of worm growth.

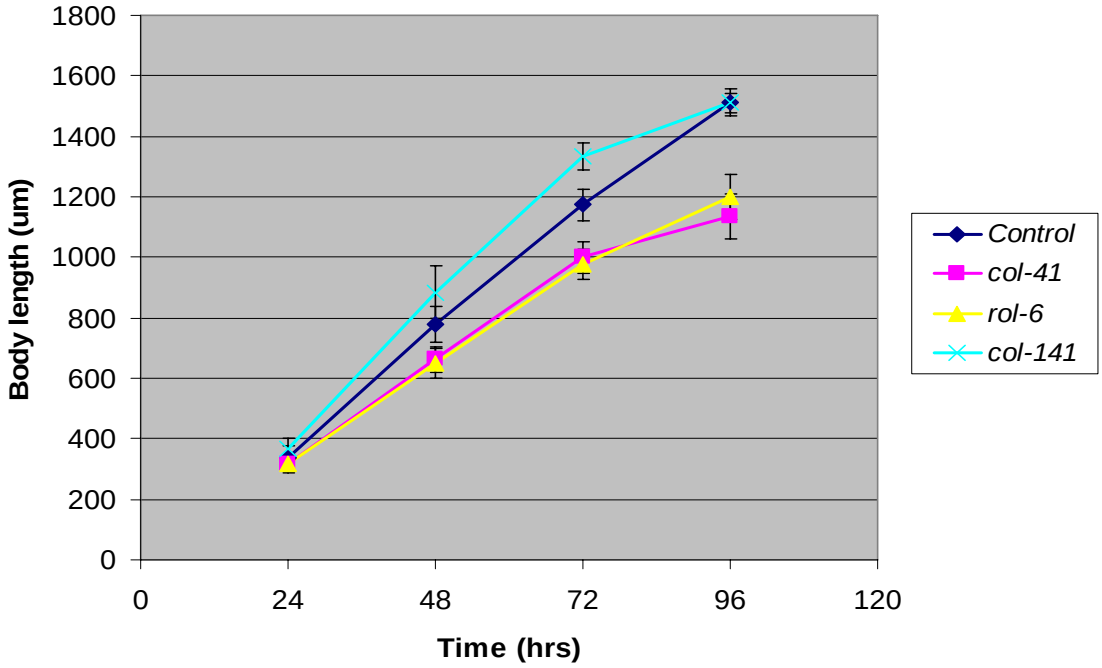


Figure 2.3 Cuticle collagen genes physiological function. Growth curve of the worm fed with *rol-6*, *col-41* and *col-141* RNAi clone.

RNAi experiments and microarray analysis demonstrate the involvement of these cuticle collagen genes in body size regulation. We next investigated where and when the collagen genes are expressed and how they are regulated by the DBL-1/BMP pathway. In order to investigate gene expression *in vivo*, I constructed transcriptional GFP reporters for *col-37*, *col-141* and *col-41*, using their upstream genomic sequences fused with GFP. However, the *col-37p::GFP* and *col-141p::GFP* reporters do not show detectable expression *in vivo*. In contrast, the *col-41p::GFP* reporter containing 1.6kb promoter region indicates a weak but detectable expression in the hypodermal syncytium (Fig 2.4), which is the critical tissue for regulation of body size. To increase the intensity of the fluorescent signal, we used a mCherry reporter containing 2X nuclear localization signals

(NLS). We expected that this construct would concentrate the fluorescent protein expression into the nucleus and, as a result, enhance the visibility for further assays.

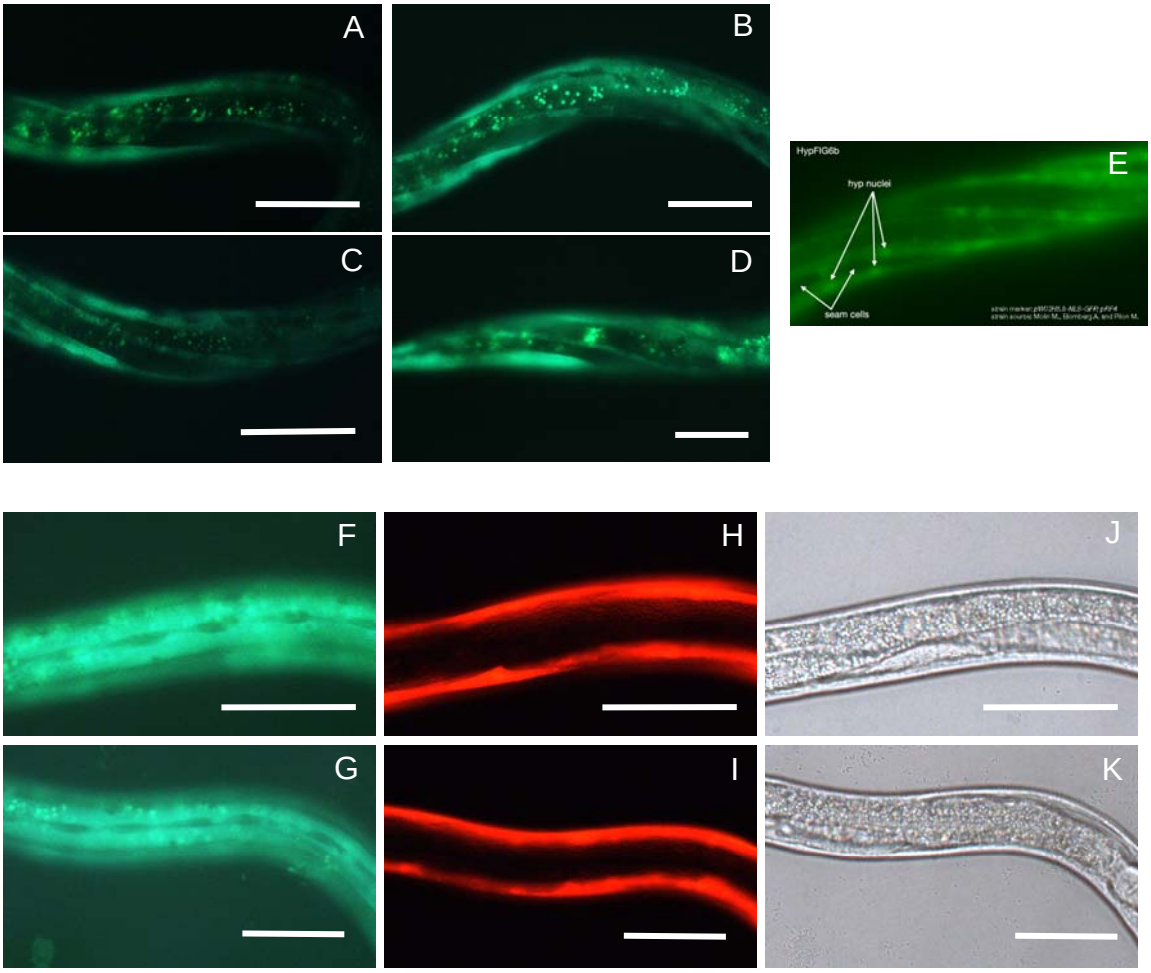
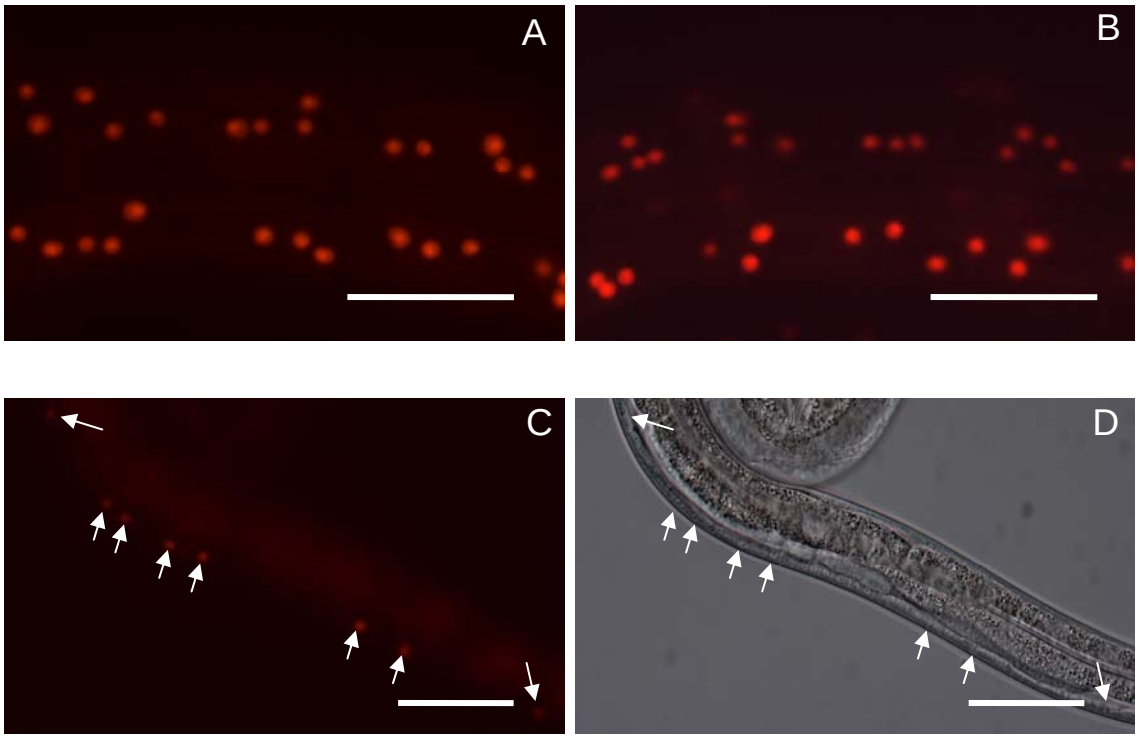


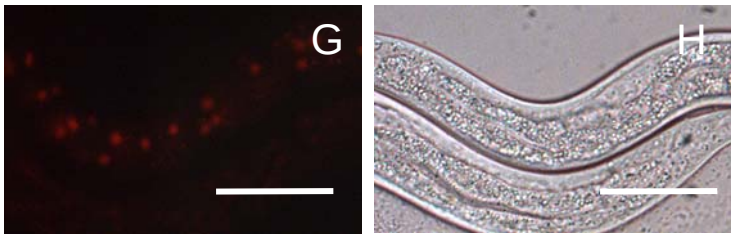
Figure 2.4 *col-41* expresses in the hypodermis. (A-D) *col-41p::GFP* expression in wild-type animals using *rol-6* as co-injection the marker. (E) Referential hypodermal GFP expression (WormAtlas). (F-K) *col-41p::GFP* expression using *myo-3p::mCherry* as the co-injection marker. (H and I) show the mCherry marker expression of the worms in F and G. (J and K) are bright field images of the worms in F and G. Scale bars: 50um.

As shown in Fig 2.5, the integrated *col-41p::2xNLS::mCherry* transgene shows the expected hypodermal nuclear expression. More importantly, this expression is stage-specific. The strongest expression level is observed in L3 to early L4 stage animals. In adults, expression of *col-41* totally disappears. Consistent with microarray and real-time PCR data, this *col-41* transgene expression is repressed in all *sma-9* mutant backgrounds (Fig. 2.5).

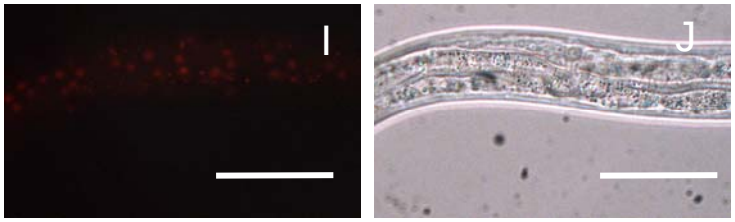
Since we are interested how *col-41* expression is regulated, I analyzed the putative promoter sequences of *col-41* in more detail. First, I compared *col-41* upstream sequences from three closely related nematodes, *C. briggsae*, *C. remanei* and *C. elegans*, I find that the -500bp to 0bp of promoter region in these three nematodes are well conserved, but not the -1.6kb to -500bp (Fig. 2.6). This conservation pattern suggests that the 500bp promoter region could be the essential regulatory element controlling *col-41* expression. Thus, I constructed mCherry reporters containing different fragments of *col-41* genomic promoter (Fig 2.7). *In vivo*, the reporter construct with the distal 1.1kb promoter fragment can not drive *col-41* expression, as compared to the full length promoter construct (data not shown). In contrast, the mCherry reporter with the proximal 500bp conserved promoter fragment results in tissue-specific (hypodermal nucleus expression) and stage-specific (expression in the early larval stages) *col-41* expression (Fig. 2.8). However, compared to transgene expression in 100% of animals containing the full length promoter, only ~60% of transgenic animals containing the integrated 500bp promoter construct showed expression (four integrated lines were analyzed). In addition, expression of the 500bp mCherry reporter construct is not repressed in any of



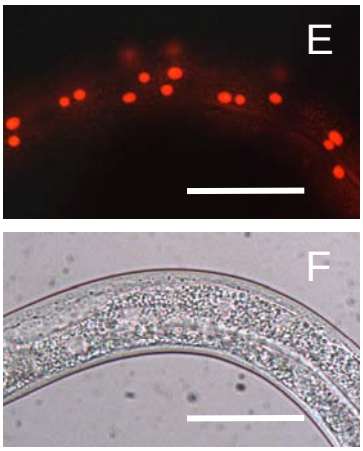
In *sma-9(qc3)* background



In *sma-9(cc604)*



In *sma-9(wk55)*



wt background

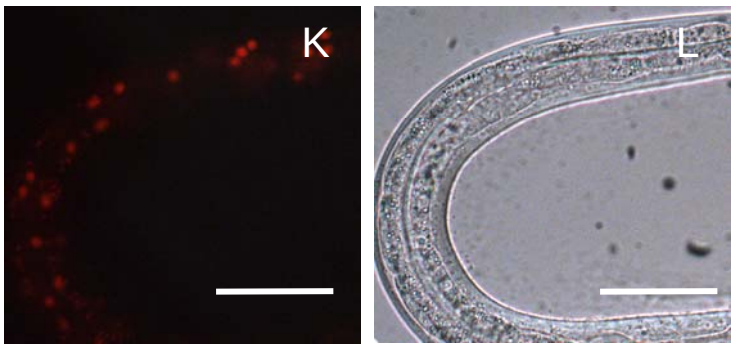


Figure 2.5 *col-41p::2xNLS::mCherry* reporter expression. (A, B), *col-41p::2xNLS::mCherry* transgene shows the expected hypodermal nuclear expression. C, Expression of mCherry in ventral hypodermal nuclei. D, Nomaski image of the worm in C. White arrows indicate the P-lineage hypodermal cell nuclei. (E-L), Consistent with microarray data and real-time RT-PCR analysis, *col-41* expression is downregulated in all three *sma-9* mutant backgrounds (G-L) compared to wildtype backgrounds (E and F). Scale bars: 50um.

```
C. elegans      CAGAGTATCAACATATGGGACTTCTGAGACATTTCCATAACTGGCCAACGACAAC--TTGGTAGTCAATTTCTTGGCGGCA--ACTTCAACATCACTAGCCGACACATTTGGAATCTGTT
C. briggsae     CAGAGTGTCGACGTATGGGACCTCGTTGATGTTTCCGTAGCTGGCCACAGAGAGT--TTGTTAGTCAATTTCTTAGCGGCA--GCGTCAACATCATCAGTGGAAACACTTGAAATCAATT
C. remanei      TACTTTCTTTATCCGCCCATTCCCCCCAGCTCTTTGCCTATTGTCATGAAATTGTGGTCGGTGGCTCTCATCTTGAAAACATGGAGCTCCCAACTACAATTGATTTATCTTGTCCTTTCT
                 *   * *  *           *        **       ** *     *      * **  *      ****     **         ** *   *   **   *  *         *

C. elegans      GAATGTAATTGTCTTGGGACTGGAAGAGTTGGGA--GGCTCTCTCGATGGCAAAGTTGTCGGAATGAGCAGA-------GGCTCTCAAGGCGTTGTTGAAGGCCTGCTTCTTGACGGCTG
C. briggsae     CAATGAAGTTGTCTTGGGACTGGAAAAGTTGGGA--AGTTCTTTCGATGGCAAAGTTGTCACTATGGGCGGA-------TGCTCTCAATGCGTTGTTGTAGGCTTGCTTCTTAACGGCTG
C. remanei      CATTATTTTTTTTCTTAG-TTTTTACAGTTGGTTTTACTTCTTTTATTCTCAAACATGTTATCCCGCACAAGCTTCCCCTGCTCACCCTACAATTTTACGGAATTTTTTATCCATGGTCC
                 * *    ** *  *  *  *   * ******       *** *   *  ****  ***      *  *          **** *    *  * **   *   *  ** *  * **

C. elegans      AAA---------GTTGAAATTAATACAAAAACCGATTAAGTCGGT-AACGTCACAGTAGTTT--------GAAATTACCATCAA------CCTAAAAATCTGCAATTAG---------TT
C. briggsae     AAAATAAAGTATGTTGGATTAAGTTGCGAAGTTGAGCGGGTATGTGAACTGTAGGAAAGTTTTAGAAAACAAAATCATCGTCAAAGATTTTTGTAAAATTAGCAATCAACACTGACTCTC
C. remanei      TACTGATGTCATTTTTTGCTAATTCCCGAATCGATTTTTCTTTTGGAGTCCCGTCTCAAATC-------CTCAATTCCCAGCTA-----TTTCACTTATCCCCGCTTCTC--------TT
                 *          **    * * *    **          *     *         *  *          ***   *  * *            **   *  *           *

C. elegans      TTGTGCTAGAAA------TATTCCATTTCGCTTAAATT-TTACGAC-------AAGGTTAGCATAGTTTAGCT--TTAGTCTAGTTTAGTTAAGTTCA---------ACTCTACCACAAG
C. briggsae     CTGCGAACGAAAATGAGTTACCGTATTTTACCGCAATTATCTCGATTCTTGTGAAAGATACCAAAAATTTGTCAACTGATAAATTTTTGTTGATATTATGATAAATAATTTCTCAGTTAA
C. remanei      ATATACATAAAATCGA---GCTTTTTCTTGTATTCAGTGTCTAATCT----TATTTATTGACAAAAAATGAGA---TGGATGGGAGGAAACAAAGTTAT--------AAATCATAACGAG
                 *       ***             * *       * * *                   *  ** *   *       *              *  * *         *         *

C. elegans      AACACAATTTATATTT-----TGGGTTTTCAGT-ATAGGCCTCTAAAAATCATATTTATACAAAAGCCCTTCTCTAAGGCGGCTCAAATTGTAAATCT----ACATA--CAAAAATGTCA
C. briggsae     CAAATTGTCAATATCTGTTACCAGGTTCAAGAT-AATTGAGCTTAAAGATTACCGTTTGGTCT--GCTTCTATTGAAGGAAATACGGTATGTAGTTTTTTGAAGGTG--CTCAGCTGTAA
C. remanei      GA-ATAATCGACTACTG----TAGGTTAAATTTTATCGGTGGTTAAAATTCTTCACTATTCAT----TTATTCACAGGGAGTCAACGTATGGAACCTCGTTGACATTTCCATAGCTGGCC
                 * *   *  *    *       ****     * *   *    ****  *      *            *    * **         ** *         *  *   *  *  **

C. elegans      AGT-----------ATTAGG--TCGTGAAAAAAGTTTT--CTCCGAAAAT-------TTTGAAAGTTTTGCCCCGGAAAGTTCTGCA------GCTTTCTGGACTATTTTATTATAGAAT
C. briggsae     GTTTGTAACTTCTGGTTAGA--AGCTAAGAATAGTTTCAATTTTGAACATATACATATTAAACAGGTTTTGGTTGGAACTTACTCCAAAATCAGCTCTCATGATCTGATTTTGGTTGGAT
C. remanei      ACGGACAACTTGCTGGTAAGCTTCTTGACGGCAGCTTCCACATCGCTAGC---------AAACAGGTTTTGGTTGGAACTTACTCCAAAATCAGCTCTCATGATCTGATTTTGGTTGGAT
                              **       * *    ** **       *                 * *  **      * *      *  *      * * **  *           * *

C. elegans      GTTG-----TAGAACATC------CTAGAAGCATATCTTTATT-TAATTTTTATTTTTTTTTATAATTTCAGAAATATCGATGTACAAAAATAGTGTGATCCCATTTTATCAAATTTATA
C. briggsae     GTTGAAAAATGAAACACTACTAGGCTGGAACCAGAAATTTACGACGACGCTCAAGTATTTTT---TTTTCCGAAAAGGCAATGTC---GAGCTATCAGAAACCATTCATGCCTTCGTATA
C. remanei      GGTTCTT--TCGATCGAG-------TAGTTGTCAGAGTGGGCTGAGGCTCTCAAGGCGTTGT-TGTGCGCTTGCTTCTTGACAGCTAAAAACGTTTTGAG---GTTAGAATTTTCTCAAA
                * *      *  * *          * *         *           * *     ** *       *         *        *    *  **     **          * *

C. elegans      CTCTAAAAAACAAA---ATTTCTTACCTTCAATATCAGCGACCTTCAAAGACTTCAATGCTCCAACTACATTCGAAACAGCTTGGTTAGCCTGGCTTCCTTCAACGACAAGGTATACTCC
C. briggsae     TTTCACAAAACTCATCAAAAACTCACATTCAATATCAGCAACCTTGAAAGATTTCAAGGCACCCACAACATTCGACACCACTTGTCCGATGTGAGCTCCATCGGCGACAATGTATGCTCC
C. remanei      ATTCCCATGAAAAA---CGTATGTACCTTCAATATCGGCAACCTTGAATGATTTCAAGGCTCCGACCACATTGGTGACGACTTGATTGGCCTGGGCACCTTCTGCAACGATGTAGACTCC
                 *    *  *   *          ** ********* ** ***** ** ** ***** ** ** ** ***** *  **  ****       **    ** **  * ** * ***  ****

C. elegans      GGCCAGTCCACTATCCGCATGAACAGCTTGGAATGCGCTGGCCGATCCATTCTGTCCGACGGCTTTTGCGATGACTCCAGTGGTTCCGGTGTTGAATCTGAAATTAATCATTTCAAGTGT
C. briggsae     GGCGAGTCCACTGTCCGAATGTACGGCTTGGTAAGCGCTGGCAGATCCATTCTCTCCGGCAGCCTTTCCAATGACTCCTGTGGCAGATGTGGAGAATCTGAAAATAA-------AAATAT
C. remanei      AGCGAGTCCGCTATCGGAGTGAACTGCTTGATAAGCGCTGGCAGATCCATGAGCTCCGGCAGCTTTAGCGAGGACACCGCTGGCAGACGAGGAGAATCTGGAAATAAA------TCATTT
                 ** ***** ** ** *  ** ** *****  * ******** *******    **** * ** **  * * *** **  ***     * *  ******* ** ***         * *

C. elegans      GAACATTGTAATTCCTTTTTTGTATTA----CTTGACTGGTGATGAATTTCCAAGTGCGGTGAGGAGAACTGCTTGGGTGGCCAATGCTTTCGTGTTGTTTCCGGCAGCTCCTTCTCCAG
C. briggsae     GGACATT-TGTCTCCATCACTGAACTAATCACTTGACTGGTGACGAATTTCCGAGGGCGGTGAGAAGAACCGCTTGTGTCGCCAAGGCCTTGGCGTTGTTTCCGGCGGCACCTTCTCCAG
C. remanei      TGTAATCTTTTGCTTTTTTTTGAATCCAT-ACTTGACTGGCGATGAGTTGCCGAGTGCTGTGAGAAGAACGGCTTGTGTTGCAAGTGCCTTCGCATTGTTTCCAGCAGCACCCTCTCCGG
                    **  *       *   ** *       ********* ** ** ** ** ** ** ***** ***** ***** ** ** *  ** ** *  ******** ** ** ** ***** *
```

```
C. elegans      CGACAATCACATGCGCGTACTTCGAATCGGCATCTCTGCGCACTTCTCCTCCCTAGGTGTCAGAAAAGGTGAGAAATGATAGTATAGATAACATAAACAGTGGAGTGGATAATGAAATTG
C. briggsae     CAACGATGACATGAGCGTACTTGGAATCGGCATCTCTGCGCACTTCTCCTCCCTAGGTGTCGGAAAAGGTGAGAAAAAATAGTATAGATAACATAAACAATGGAGTGGATAACGGAA--G
C. remanei      CGACGATGACGTGCGCGTACTTGGAATCGACGTCTCTGCGCACTTCTCCACCCTAGGTGTTGGAAAAGGTGAGAAAAGATAGTATAGATAACATAAACAGTGGAGCGGATAACGGAAG-G
                * ** ** ** ** ******** ****** * ***************** **********  **************  ********************** ***** ****** * **  *

C. elegans      AAATTTGCAAAGTTTT----TCTGTTCTATTTCCAAATCC----------AAAATGAGTGATTGCATTTTTTCGGGGAAAGAAACAATAGTAGCAAGTTTTTT--------CCTCTTC--
C. briggsae     CAATTTCAATTTTTTC----TGTGTTCTATTTCTAAGCGC----------AGA-TGAATGAGTG----------ATGAAGGACACAGTA-TAGCAAGTTTTTT--------TCCAAAAGA
C. remanei      AGATTTCAATTTTTTTGACGTGTGTTCTATTTCTAATCCCTTCTCCCCCCAGAATGAGTGATGG------TCGGGGGAAGGGGACACGAGTAGCAAGTTTTTTTTCTTTTTCCTTAGCGC
                  ****  *   ***     * ********** **   *          * * *** ***  *           *** *  ***  * **********        *

C. elegans      ATGCCA-----AAGGTATACTGAATATCCGTCGCAT-TGCGATATCTGTCATACGACTA---------------GGACCGGTAATTTTTTATCGGCG-CGTTCTTCCCGATCTGTTGCGT
C. briggsae     ATGCCGCCTATAAGGTATACAGAATTGCTGTCGCGTGTCCGACGTCGGAGATACGGCGGCTATCTGTAGTGCGGGAACCGGTAATTTTTTATCCCCAACGCTCTTTCCGATCAGTTGCGC
C. remanei      ATGCCGCC--ATAGGTATACTGAATATCCGCCGCATGTCCGACGACAT---CACGGCGATACGATACGTCCGGGGAACCGGTAATTTTTTATCG-CAACGTTCTTTCCGATCAATTGCGC
                *****       ******** ****  * * *** * * ***   *      *** *                * *****************  *  ** **** ******  *****

C. elegans      GTCTAAGTTTCT-AAGTTTTCCCAGCCTCCCC---CTTGCCCTTGCAAGTTCTTGCCGGCTTTCGAGTTTTTTC-GGAGAAAAATGTATAAAAAAGTGTTCACCAAAGCAATACACAACA
C. briggsae     CTTGAAGCTTTTGAAGTTTTTCCCACCCCCACTG-CCTGCCCTTGCAAGTTCTTGTCAGCTTTCGAGTTTTTTC-GAAGGAATATGTATAAAAAAGTGTTTGGCGAGGCTAATCACAGCA
C. remanei      CGTGAAGCCTCT-TGGTTTTCCCACCCCCATCGCTCCTGTCTTTGCAAGTTCTTATAGGATTTCGAGTTTTTTTTGAAGGGAGATGTATAAAAAAGTGTTTGGCGAAGCAAATCACAACA
                    ***  * *   ***** **  ** *  *   * ** * ************    * *************  * **  * *****************   * * ** *  **** **

C. elegans      TTGTTCTCAACG
C. briggsae     TTGTTCTCAACG
C. remanei      TTGTTCCCAACG
                ****** *****
```

Figure 2.6 Comparison between *C. elegans col-41* 1.6kb promoter region and that in two other nematodes, *C. briggsae* and *C. remanei*. * indicates identical nucleotides. They are well conserved in the -500bp to 0 bp upstream of the start codon. Alignment is performed using Clustalw.

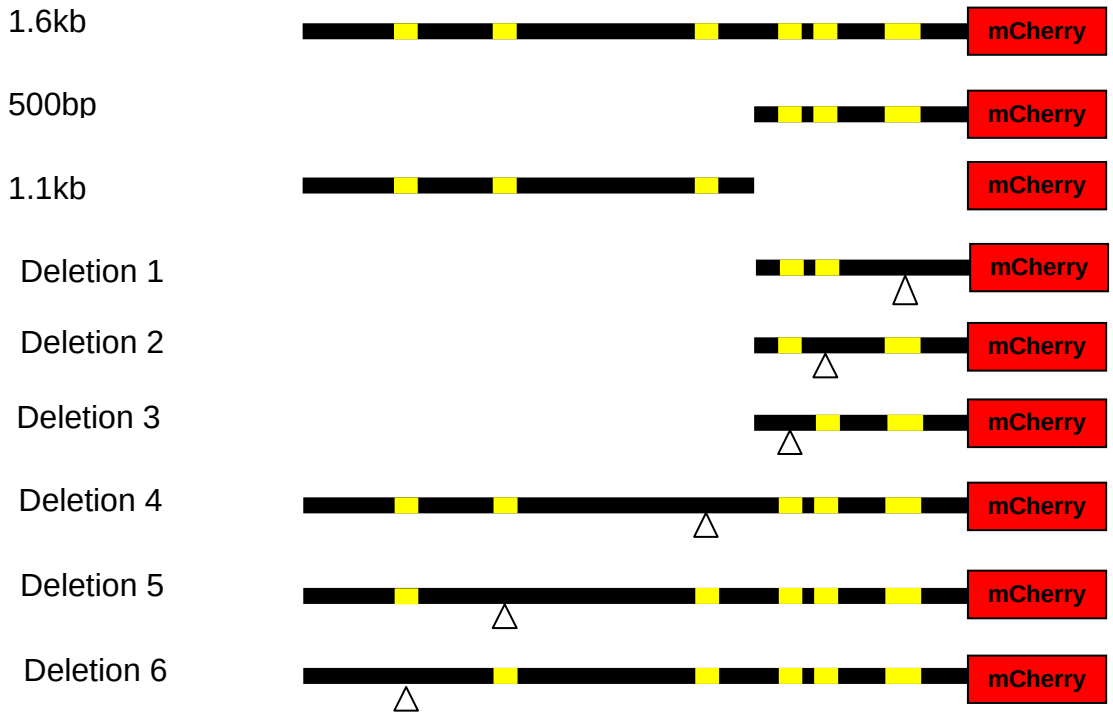


Figure 2.7 Deletion analysis of the *col-41* promoter. Red squares indicate conserved regions between three nematodes. Triangles indicate deleted regions in different deletion constructs.

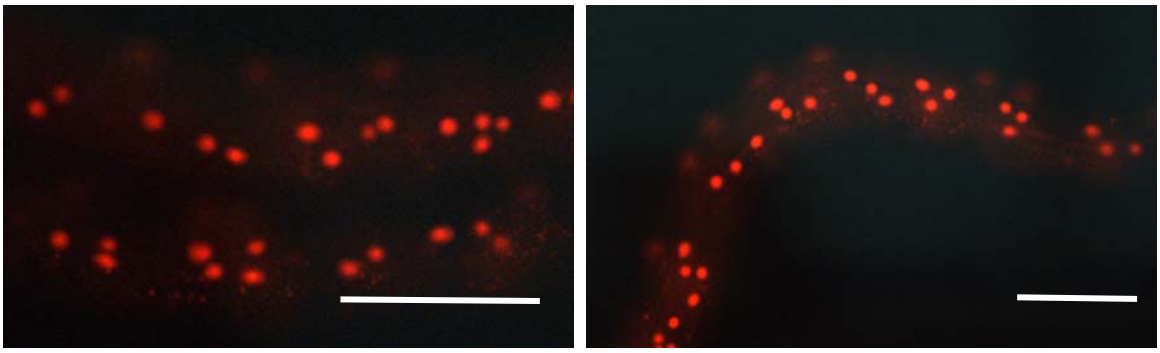


Figure 2.8 *col-41* 500bp promoter can mimic full length *col-41* promoter expression, both stage-specific and tissue specific. Scale bar: 50um**.**

the three *sma-9* mutant backgrounds (data not shown).In conclusion, the conserved 500bp promoter is necessary but not fully robust to drive normal *col-41* expression and is not responsible for the regulation of the *col-41* gene by the DBL/BMP signaling pathway.

In the 500bp promoter region, three deletion constructs were analyzed to determine which sequences are necessary for basal regulation of expression. As shown in Fig. 2.7, the three most conserved domains were deleted, respectively, to obtain three deletion reporter constructs and integrated transgenic animals carrying deletion construct were made. Deletion constructs 1 and 2 do not show detectable *col-41* expression and deletion construct 3 shows very low level of hypodermal expression (Fig. 2.9B). Since the 500bp promoter construct shows normal *col-41* expression but not response to *sma-9*, we suspect that additional *cis*-regulatory elements of the *col-41* gene are in the remaining 1.1kb region. Thus, we continued to design three deletion constructs based on the sequence conservation. As shown in Fig. 2.9, deletion constructs 4 and 5 show relatively normal col-41 expression pattern. However, deletion construct 6 shows reduced expression level, although with the normal tissue specific expression. Whether those deletion constructs are responsive to *sma-9* is still under investigation. Taken together, we believe that elements 1, 2, 3, and 6, but not elements 4 and 5, are required for *col-41* expression.

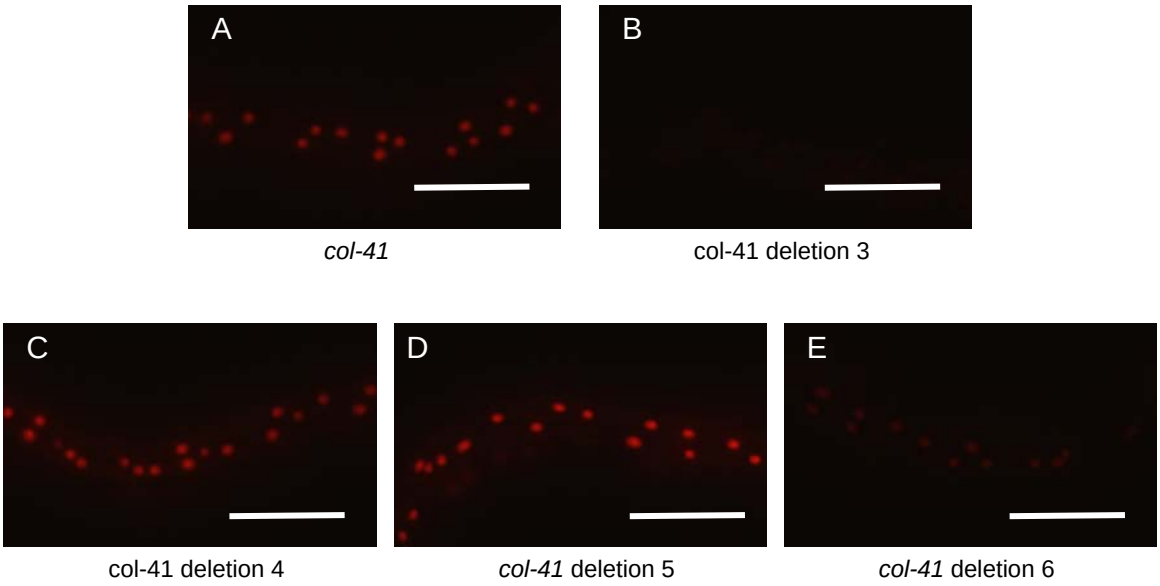


Figure 2.9 Expression of *col-41* deletion series reporter constructs in wild type. Scale bar: 50um.

Based on its expression profile and its physiological function, we conclude that *col-41* is a valid target gene of the DBL/BMP signaling pathway that mediates its role in body size regulation. In the wild type background, DBL-1 activates signaling, and through SMA-9, activates *col-41* expression to promote normal body size development. In *sma-9* mutant backgrounds, *col-41* expression is deactivated, preventing normal growth and resulting in the small body size phenotype.

Since *sma-9* is homologous to Drosophila Schnurri, we searched for the conserved Silencer Element (SE) that mediates Schnurri recruitment to regulated target genes in different organisms. However, in *col-41* promoter sequences, we do not locate the SE. We want to know whether *col-41* is a direct target gene of *sma-9*. Previously, our lab generated SMA-9 fusion constructs with SMA-9 DNA binding domain fused to

known transcriptional activator *VP-16* and repressor *Engrailed*. *sma-9* mutants carrying *en-sma-9* construct can partially rescue small body size phenotype and in reverse, *sma-9* mutants carrying *VP-16-sma-9* construct further decrease *sma-9* body size, which suggests that SMA-9 acts primarily as a repressor for body size regulation in vivo. We crossed *VP-16-sma-9* and *en-sma-9* into *col-41::mCherry* transgenic animals in the *sma-9* mutant background and examined *col-41* reporter expression. If *col-41* is a direct target gene of *sma-9*, we expected that *VP-16-sma-9* could strongly activate *col-41* reporter expression, while *en-sma-9* could further repress it. Consistent with previous results, *sma-9;Is[col-41::mCherry]* transgenic animals carrying *en-sma-9* construct partially rescue small body size phenotype while those carrying *VP-16-sma-9* construct further decrease body size. However, both constructs can activate *col-41* transgene expression to a level even higher than that of wild type animals (Fig. 2.10). These unexpected results suggest that *col-41* expression regulation could be due to a combined direct and indirect regulatory effect.

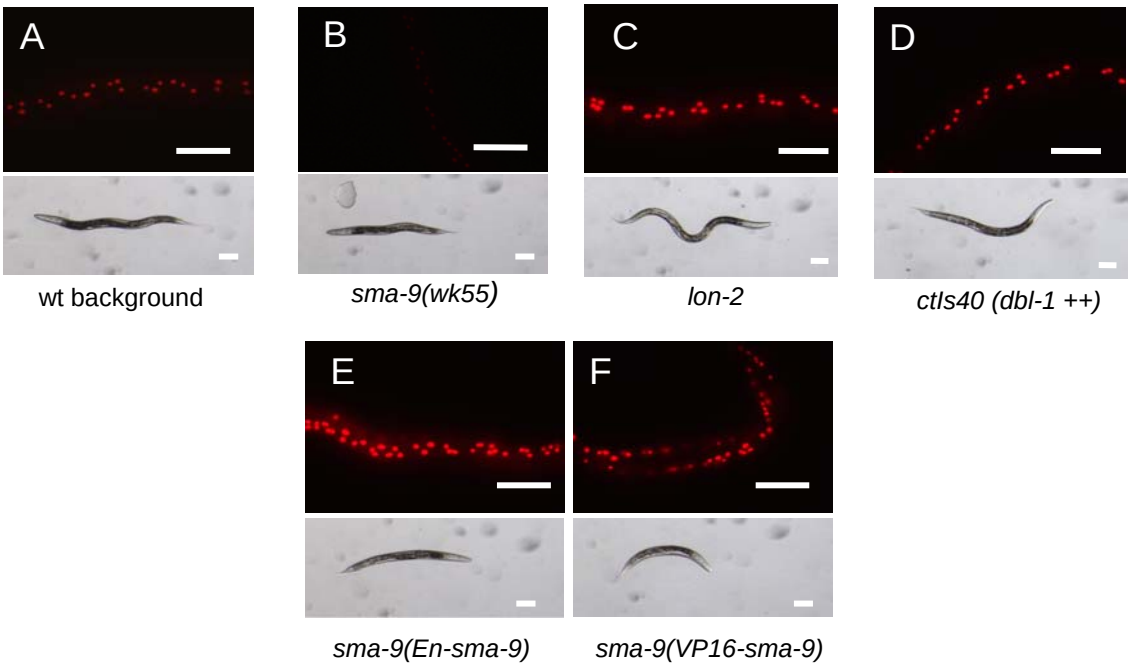


Figure 2.10 Regulation of expression of *col-41* reporters. (A, B) (C, D), *en-sma-9* and *VP16-sma-9* can both restore *col-41* expression in *sma-9* mutant background. Scale bar: 50um. (E, F) col-41 expression is upregulated in *lon-2* mutant and *dbl-1* overexpression strain. Scale bar: 50um. Lower panels of each fluorescence image are the bright field images of a representative animal of each strain at the L4 stage, indicating the proper body size phenotype reported previously (Liang et al., 2003; Liang et al., 2007; Gumienny et al., 2007; Suzuki et al., 1999). Scale bar: 100um.

The indirect regulation of *col-41* expression by *sma-9* must require other intermediate factors, such as transcription factors. Such intermediate factors should also be regulated by *sma-9*. From our previous microarray results, there are only a limited number of transcription factors regulated by *sma-9* (Table 2.2). We are able to obtain those RNAi clones form the *C. elegans* RNAi library. Unfortunately, none of those transcription factors have notable effect on the expression of *col-41* transgene in different

backgrounds (wild type, *sma-9; Is[VP-16-sma-9]* or *sma-9;Is[en-sma9]*) (data not shown).

Table 2.2 Transcription factors regulated by DBL-1/BMP signaling in *C. elegans*.

Bold: upregulated in mutant background.

Gene Identifier	Gene Function
D2030.7	**Zn-finger, C2H2 type**
F23F12.9	Basic region leucine zipper transcription factor
F46F11.2	***cey-2* - (*C. elegans* Y-box): a cold-shock/Y-box domain-containing protein; Predicted RNA-binding protein containing PIN domain and involved in translation or RNA processing; CEY-2 is predicted to function as either an RNA-binding protein involved in translation or RNA processing, or a DNA-binding protein involved in transcriptional regulation;**
K08B12.2	**Transcription factor Doublesex**
K11G9.4	*egl-46* - (EGg Laying defective): transcription factor; a TFIIA-like zinc finger protein family that affects coordinated locomotion, morphology and process formation of the touch cells, male mating efficiency, HSN cell migration, differentiation, and axonal outgrowth, serotonin production and also affects terminal divisions of the Q neuroblasts; Positively regulates; *ceh-26 , daf-19 lov-1, nlp-8 , pkd-2, egl-44*
M01E11.5	***cey-3* - (*C. elegans* Y-box) : transcription factor**
M02H5.4	**Hormone receptors; Zn-finger, C4-type steroid receptor**
R07B7.13	**Nuclear hormone receptor**
T22C8.5	**Zn-finger C2H2 type**
T27F2.4	**bZIP transcription factor**
yk220a12	***rnt-1***
yk93e2.3	***mex-5*(Muscle EXcess): CCCH-type Zn-finger protein**

II.4.3 Other potential target genes of DBL/BMP signaling pathway.

Our lab's previous microarray analysis (Liang et al., 2007) uncovered that many genes are regulated by DBL/BMP signaling pathway. Besides those potential genes discussed above, we also generated promoter fusion GFP reporter for some other target gene candidates (Table 2.3). T28A11.5 is predicted to encode a secreted cysteine rich protein and its GFP reporter does not show detectable expression in wild type animals. Warthog protein *wrt-4* and patch-related protein *ptr-22* are expressed in intestine in early larval stages and do not respond to *sma-9* regulation. However, *ptr-22* was reported to be involved in regulation of normal growth and *wrt-4* was required for normal alae formation (Zugasti et al. 2005).

Table 2.3 Summary of genes discussed in this study.

Gene	Identification	Average fold change	Background	Expression Tissue (references)
T27F2.4	Basic-leucine zipper (bZIP) transcription factor	4.03	*dbl-1* and *sma-9*	Intestine and some neurons (In this study)
T28A11.5	Predicted secreted cysteine rich protein	13.48	*dbl-1*	Undetected (In this study)
rol-6	Cuticular collagen (type IV and type XIII)	-2.15	*dbl-1*	Not Available
col-37	Cuticular collagen (type IV and type XIII)	5.73	*dbl-1*	Undetected (In this study)
col-141	Cuticular collagen (type IV and type XIII)	3.19	*dbl-1*	Undetected (In this study)
col-41	Cuticular collagen (type IV and type XIII)	-2.64	*dbl-1*	Hypodermis (In this study)
wrt-4	Warthog	-2.95	*dbl-1* and *sma-9*	Intestine (In this study and Liu et al., 2004)
ptr-22	Patched related family	3.22	*dbl-1*	Intestine (In this study)
B0218.8	C-type LECtin (clec-52)	-2.79	*dbl-1*and *sma-9*	Intestine (In this study and Murphy et al., 2003)
ins-7	Insulin-like peptide	4.23	*dbl-1*	Neurons and Intestine (Murphy et al., 2003; Liu et al., 2004)
fat-7	Acyl-CoA desaturase	-4.99	*dbl-1*	Intestine (Murphy et al., 2003)

clec-52 encodes a C-type lectin protein. Proteins that contain C-type lectin domains have a diverse range of functions including cell-cell adhesion, immune response to pathogens and apoptosis. In our microarray data, *clec-52* is downregulated in both *dbl-1* and *sma-9* mutant backgrounds. We obtained strain BC10350 from CGC, which carries *clec-52p::GFP* reporter, and crossed it into *dbl-1*, *sma-9* and *lon-2* mutant backgrounds. Surprisingly, in contrast to the microarray results, *clec-52* reporter expression is strongly upregulated in *dbl-1* mutant, downregulated in *lon-2* mutant background and has no notable change in *sma-9* mutant background (Fig. 2.11). These *in vivo* expression results suggest that *clec-52* is repressed in wild type by DBL-1/BMP signaling pathway, yet independent of *sma-9*.

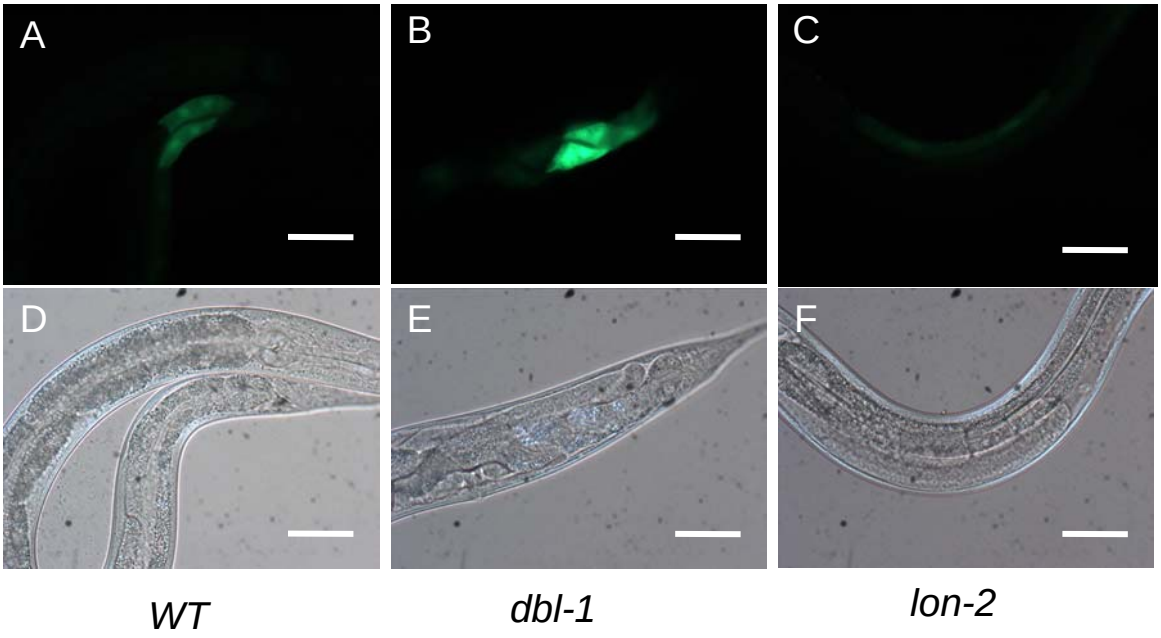


Figure 2.11 *clec-52::GFP* reporter expression. *clec-52* gene expression is highly upregulated in *dbl-1* mutant background and decreased in *lon-2* mutant background. However, its expression does not change in *sma-9* mutant background (data not shown).

D, E and F are the bright field images of the same worms in A, B and C, respectively. Scale bar: 50um.

fat-7 encodes an essential delta-9 fatty acid desaturase that is required for the synthesis of monounsaturated fatty acids. In our microarray data, *fat-7* is downregulated in both *dbl-1* and *sma-9* mutant background. We also obtained strain BC15777 from CGC, which carries *fat-7p::GFP* and crossed into *dbl-1*, *sma-9* and *lon-2* mutant backgrounds. Consistent with microarray result, *fat-7* expression is downregulated in and *sma-9* mutant background (Fig. 2.12). However, its expression does not have notable changes in *dbl-1* and *lon-2* mutant background.

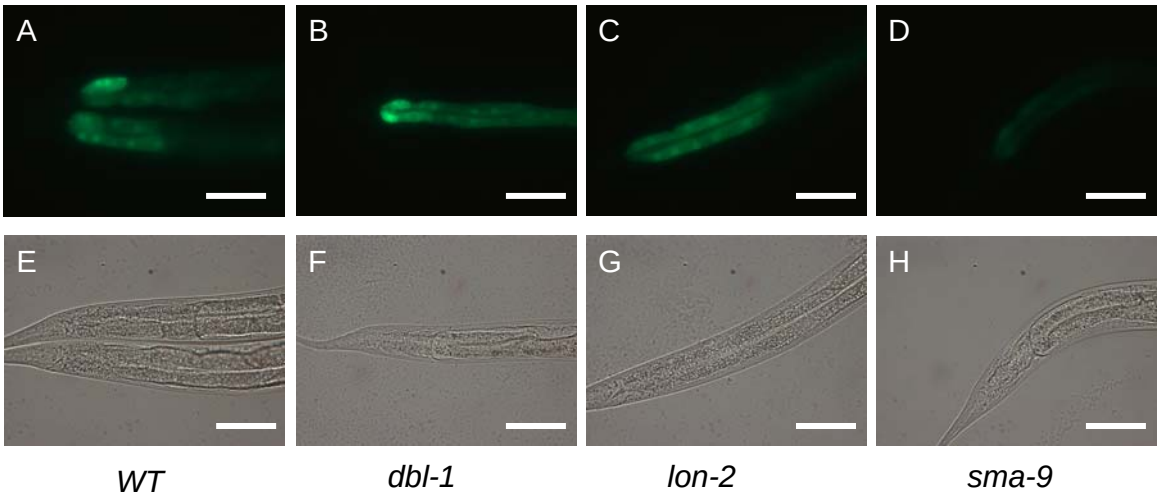


Figure 2.12 *fat-7::GFP* reporter expression. *fat-7* gene expression is greatly reduced in *sma-9* mutant background and does not have notable changes in *dbl-1* and *lon-2* mutant backgrounds. E, F, G and H are the bright field images of the same worms in A, B, C, D, respectively. Scale bar: 50um.

sod-3 encodes a iron/manganese superoxide dismutase, predicted to be mitochondrial, that might defend against oxidative stress and promote normal lifespan. In our microarray data, *sod-3* is upregulated in *sma-9* mutant. However, as shown in Fig 2.13, sod-3 GFP reporter does not show notable expression changes in different mutant backgrounds.

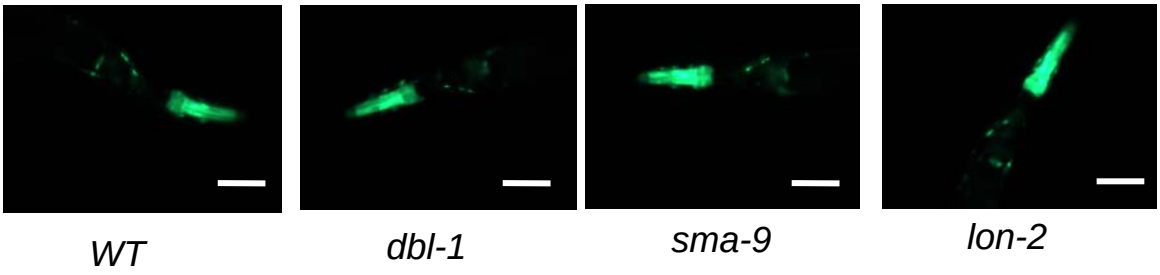


Figure 2.13 *sod-3::GFP* reporter expression. *sod-3* gene expression does not show notable changes in different mutant backgrounds. Scale bar: 50um.

II.5 Discussion

T27F2.4 gene in body size regulation

T27F2.4 gene encodes a basic leucine zipper (bZip) transcription factor homologous to human ATF-like 3. bZip transcription factors are an important transcription factor family with DNA binding activity that regulate a variety of biological processes in different organisms (Hurst. 1995). This transcription factor family includes Fos, Jun, ATFs, etc. They can form homodimers or heterodimers which bind to specific DNA sequences to regulate target gene expression. In *C. elegans*, a fos ortholog *fos-1* is required cell autonomously in the gonadal anchor cell for basement-membrane removal and subsequent anchor cell invasion of the vulval epithelium. In addition, *fos-1* activity is also required for proper vulval and uterine development, fertility, and oogenesis (Sherwood et al., 2005). Here, we provide *in vivo* evidence that T27F2.4 gene is capable of regulating body size development in response to the DBL-1/BMP pathway. T27F2.4 transcriptional reporters indicate that its expression is mainly in intestine and repressed by *sma-9*. T27F2.4 translational reporter shows similar expression pattern as the transcriptional reporter. However, it does not show significant upregulation in *sma-9* mutant background (data not shown). This could be because of a transient requirement for T27F2.4 in early larval stages and rapid turnover of T27F2.4 protein.

In Drosophila, the *sma-9* homolog Schnurri is recruited by Mad/Med complex to the SE in the *brinker* gene promoter. The consensus SE composes of a GC rich element for Mad binding and a GTCT sequence for Medea binding, which are separated by a 5bp spacer. Downstream of the T27F2.4 gene, we found a SE-like sequence, which might be

expected to be a regulatory element for T27F2.4. However, the reporter containing the SE-like element does not show differences in expression or regulation compared to the reporters containing only upstream promoter previously published from our lab. The reporter containing only upstream promoter still can responses to *sma-9* regulation (Liang et al 2003). This suggests that the SE-like fragment does not have regulatory role in T27F2.4 gene expression.

Since T27F2.4 gene is expressed in the intestine, while Smad activity is required in the hypodermis for body size regulation, it may be an indirect target of the Smad complex. In addition, the T27F2.4 potential upstream regulatory region is more than 8kb, making it an impractical tool for the analysis of cis-regulatory sequences. We therefore analyzed additional target genes to identify those expressed in the hypodermis.

Collagen genes in body size regulation

To date, existing evidence has proved the involvement of cuticle collagen genes in body size development in *C. elegans*. Some examples are *dpy-2, dpy-7, dpy-10, dpy-13, sqt-1, sqt-3, lon-2* and *lon-3* (Johnstone, 2000; Johnston et al., 1992; Kramer, 1994; Kramer and Johnson, 1993; Kramer et al., 1988; Levy et al., 1993; Nystrom et al., 2002; Suzuki et al., 2002; van der Keyl et al., 1994; von Mende et al., 1988; Gumienny et al., 2007; McMahon et al., 2003). All of these genes encode cuticular collagens. The Dpy phenotype in *dpy-5(e61)* and *dpy-11(e224)* mutants are associated with failure of the circumferential contractions in the lateral cuticle and also associated with branching or

bifurcated alae. *dpy-4, dpy-5* and *dpy-13* mutants have closely located annuli which explain longitudinal contractions in the cuticle that results in short body length when compared to wild type (Thein et al., 2003).

Our current evidence indicates that some DBL-1/BMP regulated collagen genes are also involved in body size regulation. Among them, *col-41* gene shows stage-specific expression in hypodermis. *col-37* and *col-141* transcriptional reporters do not show detectable expression, probably because of the low expression level or transgenic array silencing. We do not detect *col-41* expression in embryonic development nor in adult, which suggest *col-41* is a potential *sma-9* target gene since *sma-9* is required only in early larval development and may be dispensable later (Liang et al., 2003).

col-41 promoter analysis demonstrates that 1.6kb of upstream sequences are sufficient for expression and for regulation by *sma-9*. The proximal 500bp of upstream sequences do not contain *cis*-regulatory elements responsible for regulation by *sma-9*, yet are critical for *col-41* basal level expression. The fact that *VP-16-sma-9* and *en-sma-9* fusion constructs can both restore *col-41* expression driven by the 1.6kb promoter to a level even higher than that in wild type indicates the complexity of *col-41* gene expression regulation *in vivo*. Its expression control by *sma-9* may result from both direct and indirect effects. In order to prove its direct regulation by *sma-9*, we tested *cis*-regulatory elements in *col-41* promoter by constructing a series of deletion reporters, which are mainly based on the sequence homology of the promoters between three closely related nematodes. In order to characterize the indirect regulation, we tested

putative intermediate effectors by using RNAi to knockdown the transcription factors from our microarray analysis. Unfortunately, all the candidates we tested have no detected effect on *col-41* transgene expression in different backgrounds.

Construction of reporters for DBL-1/BMP pathway activity

In Drosophila and in vertebrates, a variety of transcriptional reporters are available to probe TGFβ signaling activity. In *C. elegans*, however, only two have been previously described. One contains an artificial enhancer composed of GTCT Smad binding sites (Tian et al., 2010); the other is an immune-response gene, *spp-9/saposin* that is expressed in the intestine (Roberts et al., 2010). We have now greatly expanded the potential toolkit of reporters for analysis of DBL-1/BMP signaling. The *col-41* gene expresses in hypodermis and is positively regulated by DBL-1/BMP signaling. The T27F2.4 gene expresses in intestine and is negatively regulated by DBL-1/BMP signaling. The *clec-52* gene expresses in intestine and is negatively regulated by DBL-1/BMP signaling independently of *sma-9*. Finally, *fat-7* expresses in intestine and is positively regulated by DBL-1/BMP signaling.

Chapter III. Alternative trans-splicing of *Caenorhabditis elegans sma-9/schnurri* generates a short transcript that provides tissue-specific function in BMP signaling

Jianghua Yin, Ling Yu, Cathy Savage-Dunn. 2010. BMC Mol Biol. 11:46.

(Contribution: JY performed the heat shock and RT-PCR experiments and participated in interpretation of data. LY performed the RNAi experiments. CSD conceived of the study, participated in its design and coordination and analyzed mutant phenotypes.)

III.1 Abstract

Background: Transcription cofactors related to *Drosophila* Schnurri facilitate the transcriptional programs regulated by BMP signaling in *C. elegans*, *Drosophila*, Xenopus, and mouse. In different systems, Schnurri homologs have been shown to act as either agonists or antagonists of Smad function, and as either positive or negative regulators of transcription. How Schnurri proteins achieve this diversity of activities is not clear. The *C. elegans sma-9/schnurri* locus undergoes alternative splicing, including an unusual *trans*-splicing event that could generate two non-overlapping shorter transcripts.

Results: We demonstrate here that the shorter transcripts are expressed *in vivo*. Furthermore, we find that one of the short transcripts plays a tissue-specific role in *sma-9* function, contributing to the patterning of male-specific sensory rays, but not to the regulation of body size. Based on previous results, we suggest that this transcript encodes a C-terminal SMA-9 isoform that may provide transcriptional activation activity, while full length isoforms may mediate transcriptional repression.

Conclusion: The alternative *trans*-splicing of *sma-9* may contribute to the diversity of functions necessary to mediate tissue-specific outputs of BMP signaling.

III.2 Background

The transforming growth factor β (TGFβ) superfamily comprises a large number of secreted peptide growth factors that have major regulatory effects on cell growth and differentiation (Roberts and Sporn, 1993; Massague and Chen, 2000). Members of this superfamily include the TGFβs, the prototypes of the superfamily; the bone morphogenetic proteins (BMPs); and other members such as activin, inhibin, and Nodal. TGFβ superfamily ligands bind to a heteromeric receptor complex at the cell surface. This complex contains two related transmembrane serine/threonine kinases, the type I and type II receptors (Wang et al., 1992; Liu et al., 1995). Signaling downstream of the receptors is mediated by the Smad proteins, which shuttle between the cytoplasm and nucleus to regulate target gene expression (Inman et al., 2002). Type I receptors directly activate receptor-regulated Smads (R-Smads) by phosphorylation at C-terminal SXS sequences (Macias-Silva et al., 1996; Kretzschmar et al., 1997). Phosphorylation of R-Smads promotes heterotrimeric complex formation with Co-Smads and accumulation in the nucleus to regulate gene transcription (Kretzschmar et al., 1997; Wu et al., 1997; Zhang et al., 1996; Inamn and Hill, 2002; Lagna et al., 1996). In mammals, five R-Smads are present: two (Smad2,3) that transduce TGFβ/activin/Nodal signals and three (Smad1,5,8) that transduce BMP signals. Strikingly, the TGFβ family ligands far outnumber the Smads available for signal transduction. Furthermore, many of these ligands are capable of eliciting diverse context-dependent responses. Thus, Smad complexes must be capable of mediating multiple diverse outcomes. It is thought that Smad complexes rely in part on transcription cofactors for appropriate regulation of target genes.

In the nematode *Caenorhabditis elegans,* the BMP-related factor DBL-1 regulates body size and male tail morphogenesis via a conserved receptor/Smad signaling pathway (Savage-Dunn, 2005). Using a genetic approach to uncover components of this pathway, we previously identified *sma-9,* a gene that is required for the body size and male tail patterning functions of DBL-1 (Liang et al., 2003). *sma-9* is also required for patterning of the mesodermal lineage, in which it acts antagonistically to the DBL-1 pathway (Foehr et al., 2006). *sma-9* is predicted to encode multiple protein products homologous to *Drosophila* Schnurri, a large zinc finger transcription cofactor that functions in Dpp/BMP signaling (Dai et al., 2000; Marty et al., 2000). Analysis of *sma-9* therefore provides the opportunity to elucidate the requirements for transcription cofactor function in an *in vivo* model system during the course of development. In addition to *Drosophila* Schnurri and *C. elegans* SMA-9/Schnurri, three vertebrate Schnurri homologs have been identified. These bind the κB-binding site and function in T cell development (Takagi et al., 2001; Oukka et al., 2002; Wu, 2002). Notably, vertebrate Schnurri homologs have more recently been demonstrated to mediate transcriptional regulation downstream of BMP and TGFβ ligands, indicating a conserved role for these family members in TGFβ signal transduction (Jin et al., 2006; Yao et al., 2006; Shukla et al., 2006).

The *sma-9* open reading frame (ORF) predicted from genomic sequence encodes a protein of 2170 aa consisting of an N-terminal Gln-rich domain (encoded by predicted exons 1-3) and a C-terminal domain containing seven Zn fingers (in predicted exons 12-22). The sequencing of cDNA clones, however, revealed at least eight different mRNA

species with alternative protein coding regions (Liang et al., 2003). Interestingly, like *sma-9*, human Shn-1 and Shn-3 genes undergo alternative splicing (Muchardt et al., 1992; Hicar et al., 2001), but the functional consequences of this processing have not been addressed. The *sma-9* cDNA clones were classified based on their potential to code for the seven C-terminal Zn fingers: class I encodes all seven Zn fingers (ZF1-7), class II encodes the first pair and the triplet of Zn fingers (ZF1-5), and class III encodes only the first pair (ZF1-2). To determine whether these domains have different functions *in vivo*, Foehr et al. created genomic/cDNA hybrid constructs and tested their ability to rescue the body size and mesodermal patterning defects of *sma-9* mutants (Foehr et al., 2006). They found that constructs encoding class I and class II C-termini were capable of rescuing both phenotypes, while the class III construct was not, suggesting that the presence of the Zn finger triplet is critical for function in body size and mesodermal lineage regulation.

Since *sma-9* is a large gene, most of the existing cDNA clones contain incomplete coding sequences that are missing the 5' end of the gene. Two cDNA clones, *yk1285a11* and *yk1237d01*, however, are notably different (Figure 3.1). Although these two cDNA clones are less than half the length of the predicted transcript, both of them have all of the hallmarks of a full-length cDNA: *trans*-spliced leader sequence, poly(A) tail, and a complete ORF. In *C. elegans*, many transcripts are processed at the 5' end by *trans*-splicing, which results in the addition of a 22-nucleotide splice leader sequence, SL1 or SL2 (Blumenthal. 2005). The mechanism of *trans*-splicing is similar to that of *cis*-splicing (intron removal), except that the splice donor sequences are provided by the SL1 and SL2 genes, rather than being contained within the context of the individual gene (Blumenthal, 2005). About half of all *C. elegans* genes are subjected to SL1 *trans*-

splicing at the 5' end. A smaller subset of *C. elegans* genes are transcribed in polycistronic operons. In an operon, the 5'-most gene generally receives the SL1 splice leader. The individual downstream genes in an operon are then separated via SL2 *trans*-splicing to the 5' ends of each of the downstream genes. SL2 *trans*-splicing is accompanied by polyadenylation to create the 3' end of the neighboring gene upstream. *yk1285a11* contains the SL1 splice leader and sequences from predicted exons 1-7. *yk1237d01* contains the SL2 splice leader and sequences from predicted exons 11-25. Thus, the structures of the *sma-9* cDNA clones *yk1285a11* and *yk1237d01* suggest that they may represent two transcripts processed from a single longer transcript by SL2 *trans*-splicing and polyadenylation as normally occurs in a *C. elegans* polycistronic operon.

Since each of these *trans*-spliced short *sma-9* transcripts was only represented by a single cDNA clone, we could not be certain that the clones were not due to cloning artifacts or a rare spurious event. Furthermore, our previous analysis did not address whether these short transcripts provide any function required for DBL-1 signal transduction. We therefore address here several remaining questions about these predicted transcripts. First, can we verify that the mRNA variants represented by *yk1285a11* and *yk1237d01* are expressed *in vivo*, rather than being artifacts produced during cDNA library construction? If so, then how is the expression of these variants regulated? Finally, do these variants show functional differences between each other and/or relative to full-length transcripts *in vivo*? In this work, we will refer to the splice variants represented by clones *yk1285a11* (5') and *yk1237d01* (3') as A11 and D01, respectively.

III.3 Results

III.3.1 Isolation of cDNAs spanning upstream and downstream regions

The *sma-9* ORF predicted from genomic sequence is encoded by 25 exons (Liang et al., 2003; Figure 3.1). Previously characterized cDNA clones, however, contained only a subset of these 25 exons and, in particular, none of the existing cDNA clones spanned a region including both the N-terminal Gln-rich domain and the C-terminal Zn finger domains. We therefore used primers in exons 1 and 25 with the potential to generate nearly full-length cDNA inserts by RT-PCR. After RT-PCR, the clones with the longest inserts were selected for sequencing. The most complete cDNA clone isolated by this approach was pCS368, in which exons 4-8 and part of exon 9 are spliced out (Figure 3.1). This transcript form is reminiscent of a form identified by the ORFeome project, in which exon 3 becomes spliced to exon 20 (Lamesch et al., 2004). These results thus support the existence of *sma-9* transcripts capable of encoding both the Gln-rich and the Zn finger domains. Additional support derives from RT-PCR data using exon 9 primers (see below).

III.3.2 Expression of *sma-9* variant transcripts

We next sought to obtain evidence that the short transcripts represented by cDNA clones *yk1285a11* and *yk1237d01* (A11 and D01) are expressed *in vivo*, rather than cDNA artifacts. To address this question, we needed an approach to distinguish these short transcripts from full-length transcripts containing the same internal sequences. Unfortunately, we were unable to detect *sma-9* mRNAs by Northern blot (data not shown). Instead, we used the unique termini of the short transcripts to design variant-specific primer pairs for RT-PCR. For A11 we employed an anchored oligo-dT-

containing reverse primer with a gene-specific internal forward primer; for D01 we used an anchored SL2 primer with a gene-specific internal reverse primer. To determine empirically whether these primer pairs are specific for A11 and D01, we performed RT-PCR using a gradient of annealing temperatures (50°C – 55°C) and analyzed the products (Figure 3.2A). In both cases, the primer pairs amplified a single band of the correct size (330 bp for A11 and 300 bp for D01), indicating that the primers specifically detect the desired transcripts. Since the higher annealing temperatures resulted in reduced yield, we used 50°C for all subsequent analyses. Based on these results, we conclude that transcripts polyadenylated downstream of exon 7 and transcripts containing the SL2 splice leader sequence upstream of exon 11 are represented in the pool of *sma-9* mRNAs *in vivo*.

The initial RT-PCR experiments were performed non-quantitatively on mixed-stage RNA preparations. We subsequently used real-time quantitative RT-PCR (qRT-PCR) to determine the relative expression levels of *sma-9* variant transcripts and whether they are developmentally or sex-specifically regulated. Since we expect full-length transcripts, but not the short transcripts, to contain sequences from exon 9, we used exon 9 internal primers to determine the expression levels of potentially full-length transcripts and calculated the abundance of the short transcripts relative to exon-9-containing transcripts. Three developmental time points were assayed: 24 hr (first larval stage – L1), 44 hr (approximately L3), and 96 hr (adults). We find that D01 transcripts accumulate in wild-type strain N2 at 45% - 85% of the level of exon-9-containing transcripts at all developmental time points examined, while A11 transcripts accumulate at approximately

10% of the level of exon-9-containing transcripts (Figure 3.2B; Table 3.1). To test for sex-specific expression, *him-5* (high incidence of males) mutant populations that contain both hermaphrodites and males were compared to wild-type N2 hermaphrodite populations. Comparing these strains, we find no evidence that the short transcripts are produced in a male- or hermaphrodite-specific manner. Thus, the A11 and D01 cDNA clones represent detectable mRNAs that are expressed during larval and adult stages when DBL-1 is active in regulating body size and male tail patterning.

Alternative splicing is regulated by *trans*-acting factors that influence the choice of splice sites. In *C. elegans*, *mec-8* encodes an RNA recognition motif-containing protein that regulates alternative splicing in the hypodermis (Spike et al., 2002; Lundquist et al., 1996). We therefore sought to determine whether MEC-8 influences that production of alternatively *trans*-spliced *sma-9* messages. We assayed A11, D01, and exon 9-containing transcript levels in *mec-8* mutants at the L3 stage by qRT-PCR. All three transcripts showed a reduction in accumulation. Relative to levels in N2 at the same stage, exon 9-containing transcripts were reduced approximately 2-fold, while the levels of the short transcripts were reduced 3- to 5-fold (Table 3.1; Figure 3.3). If *mec-8* were directly involved in processing the short transcripts from a longer precursor, then we would expect accumulation of the full-length transcripts to be increased in proportion to the decrease in short transcripts. Since our data do not demonstrate this result, we conclude that it is unlikely that MEC-8 is directly involved in this alternative *trans*-splicing reaction. Instead, the lowered accumulation of *sma-9* transcripts in *mec-8* mutants may be an indirect effect of the loss of MEC-8 activity.

III.3.3 Loss of function of *sma-9* variant transcripts

Previous work has shown that *sma-9* constructs lacking the Zn finger triplet (ZF3-5) are unable to rescue the body size and mesodermal patterning defects of a *sma-9* mutant (Foehr et al., 2006). These experiments did not test, however, for differential requirements of short vs. full-length transcripts. To test whether *sma-9* variant transcripts have specific functional roles, we performed RNAi targeting three different regions of the *sma-9* gene. In previous experiments, dsRNA was introduced by microinjection (Liang et al., 2003; Foehr et al., 2006), but we have found that RNAi by feeding results in a higher penetrance of male tail defects. We have therefore repeated these experiments using the feeding technique (Kamath and Ahringer, 2003), and extended them by targeting the internal exon 9 in addition to the 5' and 3' ends of the gene. The effectiveness of each RNAi treatment was monitored by the appearance of the characteristic small body size phenotype (data not shown).

As a component in the DBL-1 pathway, *sma-9* mutant phenotypes include small body size and male abnormal phenotypes (Liang et al., 2003). One aspect of the mail abnormal phenotype in *sma-9* is the fusion of male tail sensory rays 8 and 9, two of the nine bilaterally symmetrical pairs of male-specific sensory organs. In wild-type animals, ray 8-9 fusions are also observed, but at a lower frequency. Treatment of wild-type animals with RNAi targeting exons 1-7 caused no increase in the frequency of ray 8-9 fusions (Table 3.2), although the small body size phenotype was produced. This treatment is predicted to knockdown expression of full-length and A11 variants but not

the D01 variant. Conversely, RNAi targeting exons 21-25, predicted to knockdown the D01 and full-length variants, resulted in a two-fold increase in the frequency of ray 8-9 fusions (Table 3.2). These results are consistent with our previous findings that RNAi of the 3' end of the gene results in more severe male abnormal phenotypes than inactivation of the 5' end (Liang et al., 2003). We next compared the male tail phenotype produced by RNAi targeting the central exon 9, predicted to inhibit full-length transcripts but not the short transcripts. This treatment resulted in a male tail defect of lower penetrance than that produced by RNAi targeting the 3' end (Table 3.2). Since the most severe male tail defect occurred when D01 was targeted, these experiments suggest that D01 has a role in regulating sensory ray patterning. Full-length *sma-9* products are also predicted to function in sensory ray patterning, since the inhibition of a central exon resulted in a partial male abnormal phenotype.

As an alternative approach to RNAi, we sought to repeat this analysis using genetic mutants. We took advantage of three *sma-9* alleles containing premature termination codons at different locations in the gene (Figure 3.1; Liang et al., 2003; Foehr et al., 2006): *qc3* in predicted exon 1 (within A11), *wk55* in predicted exon 9 (between the short variants), and *cc604* in predicted exon 18 (within D01). The locations of these nonsense mutations predict that the functions of the gene products encoded by different transcripts are inactivated: in *qc3*, A11 and full-length transcripts contain a premature termination codon; in *wk55*, only full-length transcripts contain a premature termination codon; and in *cc604*, D01 and full-length transcripts contain a premature termination codon. Since additional disruptions may occur due to nonsense-mediated decay (Mango,

2001), we determined the transcript expression profile for these strains at the L3 stage (Figure 3.3; Table 3.1). The expression levels of the short transcripts relative to exon-9-containing transcripts are similar in all three strains to the wild-type ratio (Figure 3.3). In the *sma-9(wk55)* background, levels of all *sma-9* transcripts tested were reduced (Table 3.1). Although the mechanism of this reduction is unknown, it is consistent with our observations that this allele causes the most severe body size phenotype (Table 3.2; Liang et al., 2003). More importantly for the interpretation of our experiments, *sma-9(qc3)* and *sma-9(cc604)* animals express the short transcripts at levels within the normal range (Figure 3.3; Table 3.1).

The usual expectation for a series of nonsense mutations is that the earliest premature termination codon will have the most severe mutant phenotype while later termination codons may cause less severe defects if those alleles encode partially functional gene products. For the *sma-9* alleles, the opposite result is seen in the male tail. The allele encoding the latest termination codon, *cc604*, causes the most severe male tail defect, with a frequency of ray 8-9 fusions greater than any of the previously characterized alleles (Table 3.2, Liang et al., 2003). Since different investigators typically vary in their quantitation of ray fusion frequencies, we also repeated measurement of the male tail defects of *qc3* and *wk55* for direct comparison. Consistent with previous results, *qc3,* the allele containing the earliest termination codon, causes a mild male tail defect, while *wk55*, containing a termination codon in exon 9, results in an intermediate frequency of ray fusions (Table 3.2). Thus, the most severe sensory ray defect is manifested in *cc604* mutants, in which both the full-length and D01 isoforms are

disrupted. In *wk55* mutants, in which the D01 transcript is expressed at reduced levels but is not disrupted by mutation, an intermediate frequency of ray fusions occurs. In *qc3* mutants, in which wild-type levels of D01 transcripts are expressed, a mild male tail defect is seen that may be due to the premature termination of full-length transcripts. Overall, with regard to male tail patterning, these mutations form an allelic series in which their levels of activity are the opposite of that expected.

In body size, a different result was obtained with these three mutants (Table 3.2). The *sma-9(wk55)* mutant shows the smallest body size. In contrast, both *qc3* and *cc604* produce less severe body length phenotypes that are mutually indistinguishable (Table 3.2). Since *qc3* and *cc604* produce indistinguishable body size defects, this analysis does not support a specific role for either the A11 or the D01 transcript in the regulation of body size. Furthermore, the more severe body size phenotype of *wk55* mutants could be due to the lower overall level of *sma-9* expression in these mutants (Table 3.1). The results of body size analysis thus suggest that full-length isoforms, rather than A11 or D01, are critical for promoting growth. Based on our RNAi data and on the differential effects of the three nonsense alleles on body size and male tail development, we hypothesize that the D01 short transcript has a tissue-specific function in sensory ray development.

III.3.4 Rescue of *sma-9* mutant phenotypes by overexpression of D01 cDNA

Consistent with our hypothesis, we have previously demonstrated that overexpression of the D01 cDNA, but not of A11, from a heat shock promoter can

partially rescue *sma-9(wk55)* mutant male tail defects (Liang et al., 2007). In comparison, expression of *sma-9* from a cosmid genomic clone containing the entire coding region resulted in nearly complete rescue of the ray fusion defect (Liang et al., 2003). We next asked whether expression of *sma-9* short transcripts is sufficient to provide normal gene function in *qc3* and in *cc604* mutants, in which the A11 and the D01 transcript, respectively, are disrupted by premature termination codons. As before, transgenics carrying *hs-sma-9* constructs were subjected to heat shock during the L3 stage when sensory ray identities are being established (Lints and Emmons, 1999) and males were scored in adulthood. In *cc604* mutants, which contain a premature termination codon within the D01 transcript, expression of D01 resulted in partial rescue of the male tail defect (Figure 3.4A). In contrast, overexpression of A11 in *cc604* caused a slight reduction in frequency of sensory ray fusions, but this effect was not statistically significant (p=0.07). In *qc3* mutants, in which the A11 transcript contains a premature termination codon, we were surprised to find that overexpression of A11 produced no change in male tail phenotype, whereas overexpression of D01 resulted in significant rescue of the male tail defect (Figure 3.4A). Based on the rescue of male tail defects by D01 in both *qc3* and *wk55* mutants, we conclude that increased expression of D01 may be able to compensate partially for defective full-length *sma-9* transcripts.

Finally, we looked for evidence of A11 or D01 function in body size regulation. Since heat shock of transgenics in the L3 stage produce no evident changes in body size (data not shown), we performed heat shock earlier in development, in L1 animals, and measured body length in adulthood. In contrast to the rescue of male abnormal

phenotypes by D01, heat shock of A11 and D01 transgenics led either to a reduction in body size or to no change in body size (Figure 3.4B). These results suggest that A11 and D01 play a minor role, if any, in the regulation of body size.

III.4 Discussion

We have shown here that the *C. elegans sma-9/schnurri* locus undergoes a novel *trans*-splicing process to generate two shorter transcripts, one of which has a tissue-specific function in DBL-1/BMP signaling. First, we have demonstrated that an SL2 *trans*-spliced, truncated cDNA form, *yk1237d01* (D01), represents an actual mRNA that is detectable by RT-PCR. Second, we have verified that longer transcripts containing both the 5' and 3' ends of the gene are also detectable by RT-PCR. Third, loss-of-function experiments (RNAi and premature termination mutants) demonstrate that the short transcript D01 is necessary but not sufficient for male tail patterning, since depletion of both D01 and full-length transcripts causes a more severe male tail patterning defect than depletion of full-length transcripts alone. Fourth, gain-of-function experiments (heat shock–induced overexpression) indicate that the short transcript D01 can partially rescue male tail but not body size defects of *sma-9* mutants. In contrast to the results on D01, we find for the *yk1285a11* (A11) transcript little or no evidence of an *in vivo* function, as well as low abundance throughout development. Why then is the A11 transcript detected at all? One possibility is that it is generated as a byproduct of the *trans*-splicing mechanism that generates D01.

Since we have shown that full-length transcripts and D01 transcripts are each necessary but not sufficient for male tail patterning, we must consider what differences in molecular function are found in the respective gene products. In other organisms, Schnurri homologs have been shown to act either as transcriptional activators (Jin et al., 2006; Yao et al., 2006; Torres-Vazquez et al., 2001) or as transcriptional repressors

(Marty et al., 2000; Pyrowolakis et al., 2004) mediating BMP-dependent transcription. Strikingly, Yao et al. (Yao et al., 2006) have shown by swapping *Drosophila* and Xenopus Schnurri homologs that transcriptional activation or repression activities may be context-dependent rather than an intrinsic property of the protein. In *C. elegans,* we have previously analyzed intrinsic transcriptional activities of SMA-9 protein domains using a heterologous transcription assay (Liang et al., 2007). In these experiments, the acidic residue-rich domain (ARD) in exon 14 exhibited transcriptional activator activity, while the N-terminal region of SMA-9 displayed intrinsic transcriptional repressor activity (Liang et al., 2007). Full-length *sma-9* transcripts should encode protein products containing both of these transcriptional domains, while D01 transcripts encode products containing solely the transcriptional activation domain (Figure 3.1). In addition, class I and II full-length transcripts and the D01 short transcript, but not the A11 transcript, encode the Zn finger triplet domain shown to be critical in the regulation of body size and mesodermal patterning (Foehr et al., 2006). In *Drosophila,* this domain is also critical in mediating Dpp-responsive target gene regulation (Muller et al., 2003). We therefore hypothesize that full-length SMA-9 isoforms function as transcriptional repressors or activators depending on context, while D01-encoded isoforms may be obligate transcriptional activators. This transcriptional activation activity may be necessary for the robust activation of *sma-9* target genes in sensory ray development, but may be dispensable in body size regulation. Consistent with this hypothesis, a SMA-9 Zn finger domain fusion with a known transcriptional activator, VP16, can partially rescue the male tail defects of *sma-9* mutants, but acts as a dominant-negative construct in body size

regulation (Liang et al., 2007). Thus, the *in vivo* activity of the SMA-9::VP16 fusion is strikingly similar to that of the D01 isoform.

A variety of genetic mechanisms exist for generating diverse functions from a single genetic locus, including those that affect transcript structure such as alternative splicing, alternative transcriptional start sites, and alternative polyadenylation. One prominent example of alternative splicing is the regulation of sex determination in *Drosophila* by Sex-lethal (Black, 2003). This system also uses developmentally regulated alternative promoters. In *C. elegans*, extensive use of alternative promoters leading to different protein isoforms has been documented (Choi and Newman, 2006). Alternative polyadenylation can also lead to the production of truncated protein variants, such as in the *daf-4* BMP receptor gene in *C. elegans*, in which alternative polyadenylation leads to the production of a secreted negative inhibitor of DAF-4 signaling (Gunther and Riddle, 1991). Both *C. elegans sma-9* (Liang et al., 2003) and the human homologs Shn-1 and Shn-3 (Muchardt et al., 1992; Hicar et al., 2001) display alternative splicing that is predicted to result in a variety of protein isoforms. For most of these isoforms, however, no specific *in vivo* function has been described. We have now demonstrated a functional role for the D01-encoded isoform in sensory ray development but not in body size regulation. To our knowledge, alternative *trans*-splicing as we have characterized it for *sma-9* has not previously been reported for other loci in *C. elegans*. The unique feature of this *trans*-splicing event is the choice between processing the transcript via *cis*- or *trans*-splicing at the splice acceptor site of predicted exon 11. In an organism without *trans*-splicing, a similar truncation could be achieved via use of a downstream transcriptional

start site or via proteolytic cleavage. We speculate that since the *trans*-splicing machinery exists in *C. elegans*, it was available to be recruited for this unusual role in the *sma-9* locus.

III.5 Methods

C. elegans strains and culture

Nematodes 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 I, *mec-8(e398)*; LG V, *him-5(e1490)*; LG X, *sma-9(qc3, wk55, cc604)*.

RT-PCR

Real time RT-PCR was performed on a LightCycler 2.0 (Roche). Data analysis software used was LightCycler software version 4.0. Animals were collected at desired time points and total RNA extracted by Trizol as described (Liang et al., 2003). SuperScriptTM III platinum two-step qRT-PCR kit with SYBR Green (Invitrogen) was used to perform RT-PCR. Actin gene *act-1* was used as standard control. A primer in an intron of the *act-1* gene was used to confirm the absence of genomic DNA in RNA preps.

To generate cDNA clones, RNA was extracted from mixed-stage N2 animals. Two-step RT-PCR was performed using SuperScript III First-Strand Synthesis System (Invitrogen). Primer sequences for all RT-PCR experiments are available upon request.

RNAi

RNAi feeding method was performed as described (Kamath and Ahringer, 2003) after cloning *sma-9* fragments into RNAi feeding vector pPD129.36 (a gift from Dr. A. Fire). To score sensory ray patterning, only adult males displaying a small phenotype

were collected and scored to ensure that the animals analyzed had undergone a knockdown of *sma-9* activity.

Heat-shock clone constructs and generation of transgenic animals

Two *sma-9* splice variants *yk1285a11* and *yk1237d01* (GenBank Accession Numbers: AY390537 and AY390550) were identified previously (Liang et al., 2003), which were SL1- and SL2- *trans*-spliced cDNA clones that appeared to represent shorter but complete mRNAs. Heat-shock vector pPD49.83 was a kind gift from Dr. J. Liu. *yk1285a11* and *yk1237d01* were cloned into pPD49.83 at Nhe I and Kpn I sites. Transgenic nematodes were generated by microinjection of constructs (10ng/ul) into the gonadal syncytia of him-5 hermaphrodites, with *rol-6* (100ng/ul) as a marker (Mello et al., 1991). Mutant strains carrying transgenic constructs were generated by appropriate crosses between the mutant strains and transgenic lines.

Heat-shock experiments

Gravid hermaphrodites were washed off plates with M9 buffer and collected in an Eppendorf tube. Hermaphrodites were destroyed with sodium hypochlorite solution, leaving eggs that were collected and inoculated onto OP50-seeded plates. For body size measurement, 24 hours after egg collection, worms were collected and transferred to a siliconized Eppendorf tube containing 100ul of M9 buffer (Brenner, 1974), which was placed in a circulating water bath under the heat-shock conditions specified. After heat-shock, worms were recovered, placed at 25°C on OP50-seeded plates and allowed to develop to adulthood. The length of individual worms was measured using Image-Pro

Express software (Sigma) with Nomarski optics. For male tail ray examination, worms were collected 72 hours after egg collection and transferred to a siliconized Eppendorf tube containing 100ul of M9 buffer (Brenner, 1974), which was placed in a circulating water bath under the heat-shock conditions specified. After heat-shock, worms were recovered, placed at 25°C on OP50-seeded plates and allowed to develop to adulthood and then scored for ray patterning.

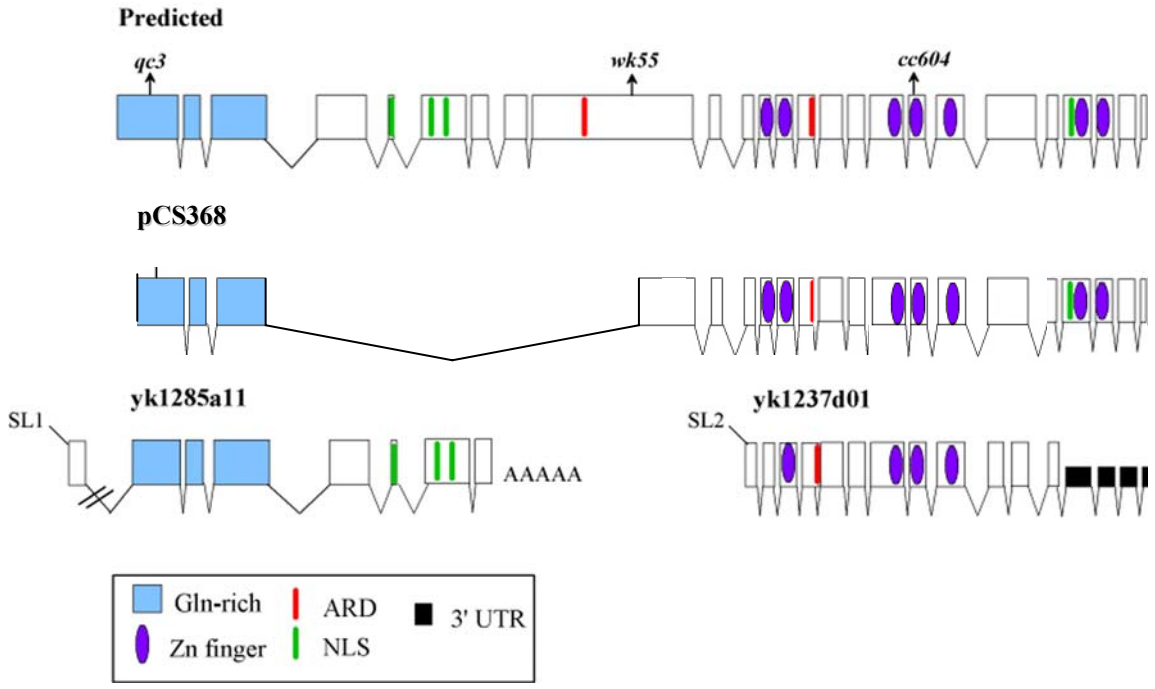


Figure 3.1 Structure of *sma-9* transcript variants. The predicted *sma-9* intron-exon structure is shown above. Arrows mark positions of nonsense codons in three *sma-9* alleles. pCS386 is the most complete cDNA variant isolated to date. The two short variants represented by cDNA clones *yk1285a11* (A11) and *yk1237d01* (D01) are shown below. ARD: acidic residue-rich domain; NLS: nuclear localization signal.

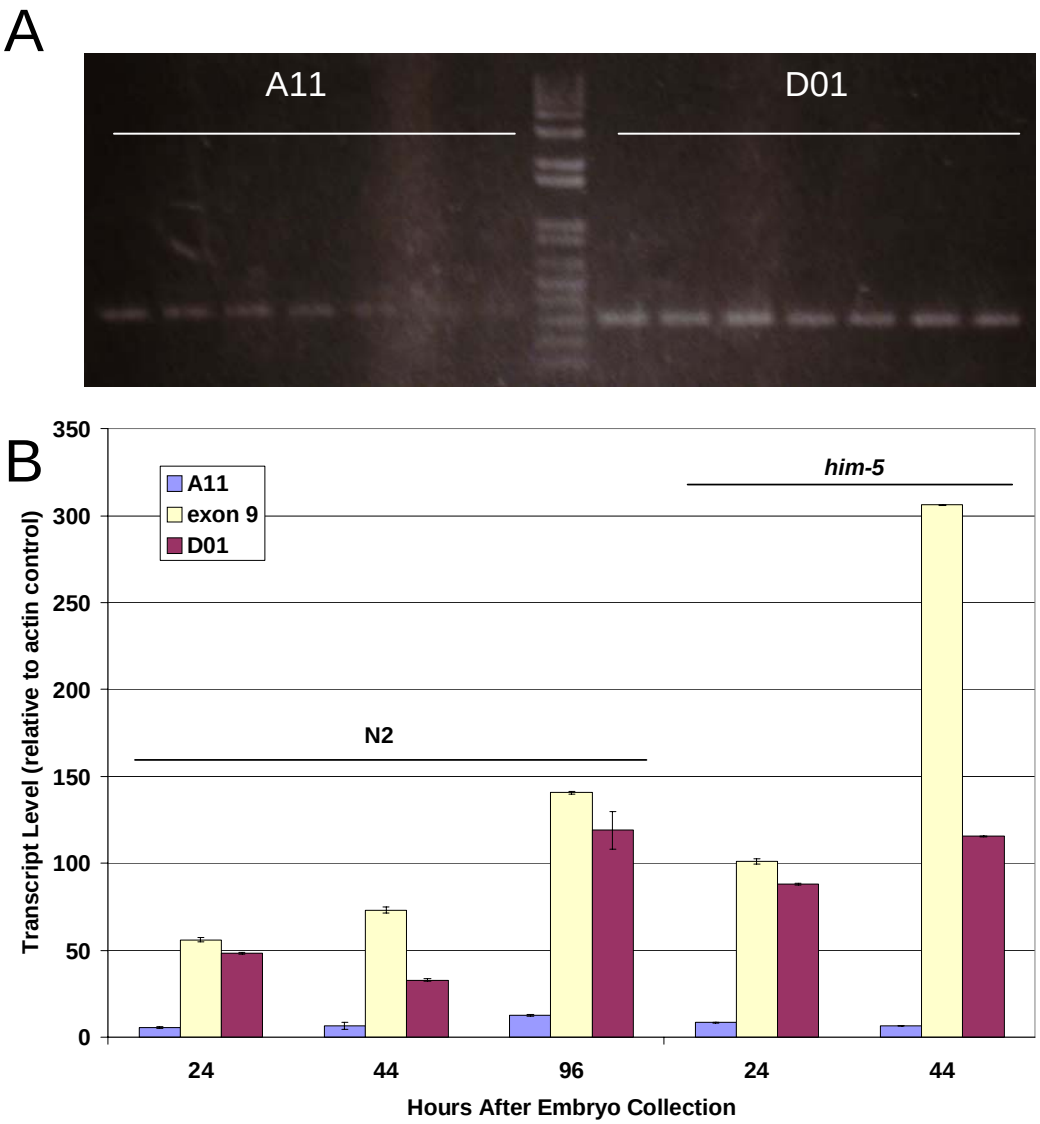


Figure 3.2 Expression of *sma-9* transcript variants. A. Detection of A11 and D01 transcripts using variant-specific RT-PCR over a gradient of annealing temperatures (left to right: 50°C – 55°C). The expected size of the PCR product is 330 bp for A11 and 300 bp for D01. **B.** Developmental profile of *sma-9* transcript expression levels, as determined by qRT-PCR. Data are shown as transcript abundance relative to *act-1* actin gene control, and error bars show standard deviation. See Table 3.1 for quantitation of transcript abundance relative to exon-9-containing control.

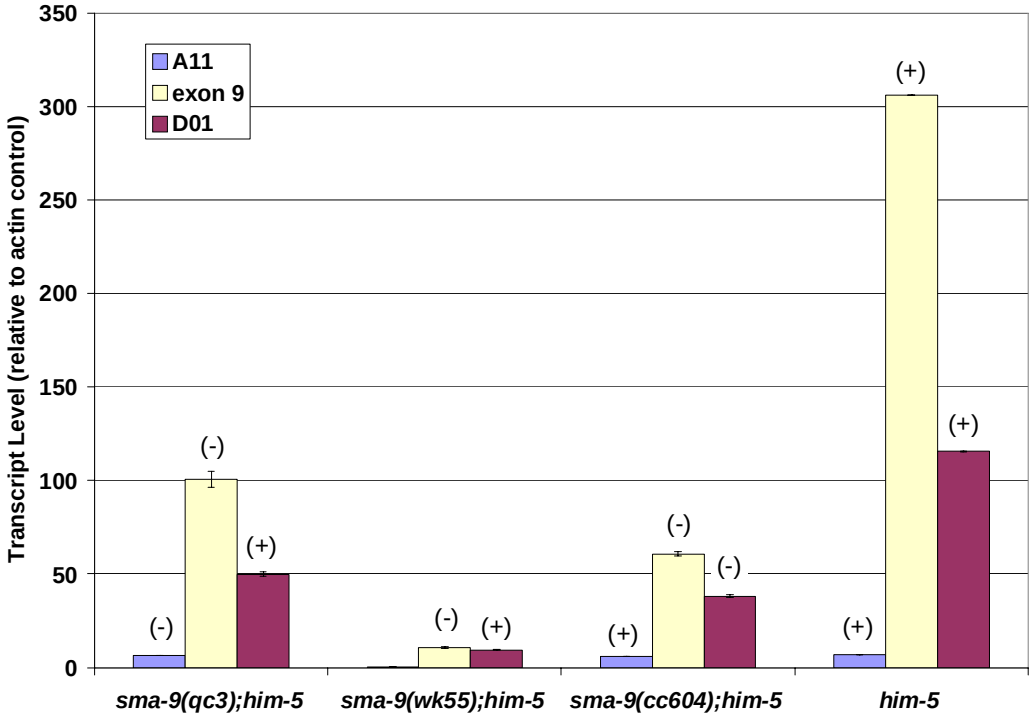


Figure 3.3 Expression levels of *sma-9* transcripts in mutant backgrounds. Expression levels of *sma-9* transcripts as determined by qRT-PCR. Animals were collected at the L3 stage. Data are shown as transcript abundance relative to *act-1* actin gene control, and error bars show standard deviation. See Table 3.1 for quantitation of transcript abundance relative to exon-9-containing transcripts. (-): transcript harbors a premature termination codon. (+): transcript sequence is normal.

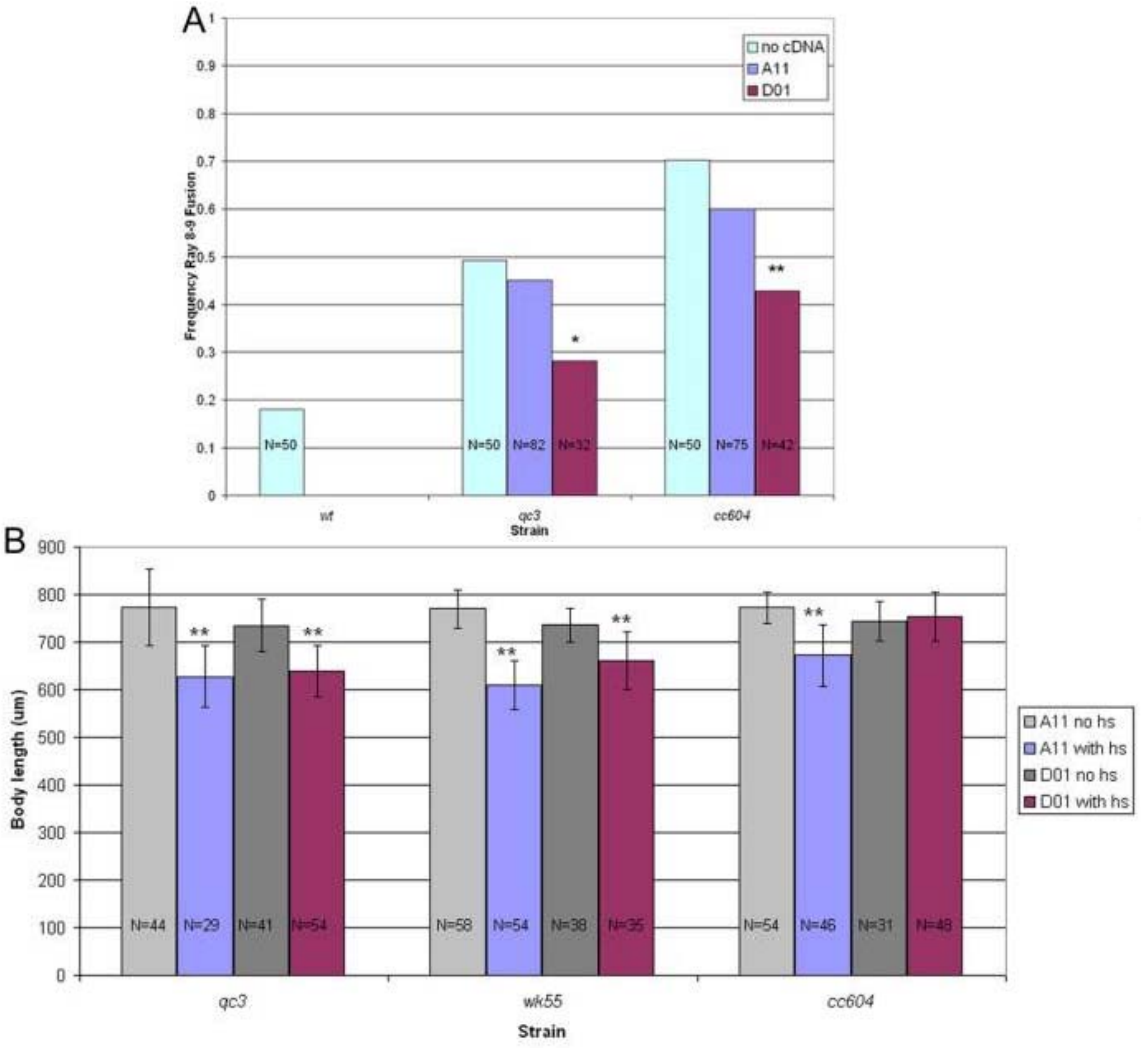


Figure 3.4 Phenotypic rescue by overexpression of *sma-9* short transcripts. A. Male tail phenotypes are given as frequency of fusion of rays 8 and 9 per male tail side scored. p values were calculated using a Student's t-test. *: significant difference from no cDNA control ($p<0.05$). **: highly significant difference from no cDNA control ($p<0.01$). The p values for experiments with A11 cDNA overexpression showed no significant difference from controls (p=0.59 for *qc3*; p=0.16 for *cc604*). All strains contain *him-5(e1490)* to increase the proportion of males. **B.** Body size phenotypes. Mean body length at 96 hours after embryo collection (adult stage) is shown. Error bars show standard deviation. Note that these strains are transgenics expressing the *rol-6* marker that influences body length,

so all comparisons are made between heat-shocked animals and no heat shock control rather than using a nontransgenic control. p values were calculated using a Student's t-test. **: highly significant difference from no heat shock control ($p<0.01$).

Table 3.1 Expression levels of *sma-9* transcript variants determined by qRT-PCR.

Strain	Developmental Timepoint	A11 Transcript Level	D01 Transcript Level	Exon 9 Transcript Level
N2	L1	5.66 ± 0.59	48.27 ± 0.66	55.90 ± 1.19
N2	L3	6.50 ± 1.89	32.92 ± 0.64	73.17 ± 1.88
N2	adult	12.61 ± 0.60	118.96 ± 11.00	140.65 ± 0.74
him-5	L1	8.31 ± 0.28	87.90 ± 0.38	101.02 ± 1.70
him-5	L3	6.68 ± 0.00	115.51 ± 0.35	306.17 ± 0.18
mec-8	L3	1.53 ± 0.16	10.63 ± 0.57	38.55 ± 9.85
sma-9(qc3)	L3	6.25 ± 0.01	49.96 ± 1.31	100.84 ± 4.24
sma-9(wk55)	L3	0.19 ± 0.01	9.39 ± 0.14	10.61 ± 0.43
sma-9(cc604)	L3	5.70 ± 0.03	38.03 ± 0.76	60.66 ± 1.15

Mean expression levels are give as percent of the expression level of the *act-1* actin gene control, plus or minus standard deviation.

Table 3.2 *sma-9* loss of function phenotypes in male tail patterning and body size.

RNAi (Exon) or Mutant	Forms Predicted to be Blocked	Frequency of Ray 8-9 Fusions[1] (n)	Body Length[2] (n)
none	none	18% (50)	ND
1-7	full-length, A11	20% (50)	ND
9	full-length only	29% (100)	ND
21-25	full-length, D01	36% (50)	ND
sma-9(qc3)	full-length, A11	50% (50)	692 ± 85 (43)
sma-9(wk55)	full-length only	66% (50)	566 ± 34 (31)
sma-9(cc604)	full-length, D01	72% (50)	716 ± 61 (30)

[1]All male tail phenotypes are scored with *him-5(e1490)* in the background to increase the frequency of males.

[2]Body length is given in μm ± standard deviation at 96 hours after embryo collection.

ND: not determined.

Chapter IV. Conclusion

Drosophila Schnurri (Shn) is a large zinc finger DNA-binding transcription factor and was one of the first partners identified for Bmp-specific R-Smads. Soon, Shn homologs were found in vertebrates (Shn-1, Shn-2 and Shn-3) and *C. elegans* (SMA-9). In Drosophila, the function of Schnurri (Shn) has been well characterized, but how the function of Drosophila Schnurri relates to that of its homologs from other species has not been as extensively investigated. First, Shn has been shown to act as either a transcriptional repressor or as a transcriptional activator. Shn mediates the repression of the *brinker* gene and in turn derepresses a large set of *brinker* target genes. In contrast, Shn can also directly activate gene expression in response to Dpp signaling, as in the case of *Ultrabithorax* (Ubx). *sma-9* in *C. elegans* is predicted to encode a large zinc finger protein homologous to Drosophila Shn. We have shown that one *sma-9* splicing variants D01 can partially rescue male tail defect in *sma-9* mutants, which strikingly mimic the *VP-16-sma-9* function (Liang et al. 2007). Combined with the fact that *enR-sma-9* could partially rescue body size defect in *sma-9* mutants, we suggest that SMA-9 also acts as both transcriptional repressor and activator in regulating different Sma/Mab signaling outcomes.

Second, the mechanism of Shn recruitment to the Brinker gene for repression has also been well studied. It is recruited to a defined element (consist of a Co-Smad binding site 5'-GTCT-3' and a R-Smad binding site 5'-GCCGNC-3' separated by 5bp spacer) on the *brinker* gene promoter by the Mad/Med complex, repressing *brinker* expression. In addition to *brinker* gene repression, the Shn/Mad/Med complex is also found to be

essential for repression of *bam* and *gsb* genes. In vertebrates, Schnurri homolog Shn-2 is found to be recruited by Smad1/4 to three 5'-AGAC-3' SBE (Smad binding element) sites in the upsteam of PPARγ2 promoter in response to BMP. In *Xenopus*, Shn-1 could stimulate response to BMP signaling through the Bmp response element (BRE), which is similar to the defined element on *brinker* gene promoter. These conserved regulatory mechanisms raise the question that in *C. elegans*, is *sma-9* also recruited by Smad complex to a similar sequence?

In Drosophila, a few DPP/BMP signaling target genes have been identified. In this study, we provide evidence of the potential target genes in DBL-1/BMP signaling and their involvement in body size regulation. In particular, we find that T27F2.4 negatively and *col-41* positively regulate body size development. In addition, we generated reporter constructs for target genes that greatly expand the potential toolkit for future analysis of DBL-1/BMP signaling.

In *C. elegans*, DBL-1/BMP signaling activity is required for body size regulation. Previous studies have shown that DBL-1/BMP signaling to the hypodermis is necessary and sufficient for normal body size. However, little is known about the target genes of DBL-1/BMP signaling and how they contribute to body size regulation. During our verification of DBL-1/BMP signaling target genes, we only found one gene, *col-41*, which has detectable expression in hypodermis. Many potential target genes show expression in intestine and neurons. Possibly, the transcriptional reporters we constructed do not represent the complete gene expression profile. We still believe that the hypodermis is the critical tissue for body size development. However, we should also pay more attention to those genes expressed in other tissues, which could have indirect effects

on the regulation of body size.

Appendix

Transcriptional Repressor and Activator Activities of SMA-9 Contribute Differentially to BMP-Related Signaling Outputs

Jun Liang[1], Ling Yu, Jianghua Yin, and Cathy Savage-Dunn

Department of Biology, Queens College, and Biochemistry PhD Program, the Graduate School and University Center, the City University of New York, Flushing, NY 11367

[1]Current address: Department of Developmental Hematopoiesis, Cell Biology Program, Memorial Sloan-Kettering Cancer Institute, New York, NY 10021

Corresponding author: Cathy Savage-Dunn, Department of Biology, Queens College, CUNY, 65-30 Kissena Blvd., Flushing, NY 11367, Tel: 718-997-3440; Fax: 718-997-3445; E-mail: cathy.savagedunn@qc.cuny.edu

Abstract

In the nematode *Caenorhabditis elegans,* the BMP-related growth factor DBL-1 regulates body size and male tail morphogenesis via a conserved receptor/Smad signaling pathway. Smads are transcription factors, but rely on transcription cofactors for appropriate regulation of target genes in response to TGF-β- and BMP-related signals. In the DBL-1 pathway, *sma-9* encodes multiple zinc finger transcription factors homologous to *Drosophila* Schnurri, which functions in Dpp/BMP signaling. We have studied the molecular functions of SMA-9 as a model for transcription cofactor-dependent regulation of gene expression. Using SMA-9 fusions to known transcriptional activators and repressors, we demonstrate that SMA-9 acts primarily as a transcriptional repressor in body size regulation in vivo. In contrast, both activator and repressor functions contribute to male tail patterning. We further show that different SMA-9 regions have intrinsic repressor and activator activities using a yeast transcription assay. We use microarray analysis to identify transcriptional target genes in body size regulation. Consistent with the importance of repression in mediating body size regulation, we find more repressed genes than activated genes in this pool. Finally, we identify five transcriptional targets with body size and/or male tail patterning phenotypes, including transcription factors related to Runx and fos and signaling molecules related to hedgehog and patched. Our results thus suggest that SMA-9 products function differentially as transcriptional repressors and activators in DBL-1/BMP pathway regulated body size and male tail morphogenesis.

Keywords: *sma-9;* BMP; Schnurri; transcriptional repressor; transcriptional activator; DNA microarray; gene regulation; body size

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. Smads are signal transducers for TGF-β (transforming growth factor-β) superfamily ligands that function by regulating gene transcription. Since DNA binding sites for Smads show low complexity and affinity, transcriptional cofactors are often required for target gene regulation (Shi et al., 1998; Massague, 1998; Derynck and Zhang, 2003). Smad complexes activate or repress gene transcription depending on whether they associate with transcriptional activators or repressors. We have previously identified a gene, *sma-9*, that encodes a putative transcription cofactor in the DBL-1 signaling pathway in the nematode *Caenorhabditis 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 (Liang et al., 2003). Furthermore, SMA-9 is homologous to *Drosophila* Schnurri (Shn), which has been proposed to function as a transcription factor in Dpp signaling. Three vertebrate homologs of Schnurri have also been identified. These bind the κB-binding site and function in T cell development (Takagi et al., 2001; Oukka et al., 2002; Wu, 2002). Like *sma-9*, human Shn-1 and Shn-3 genes undergo alternative splicing (Muchardt et al., 1992; Hicar et al., 2001). Notably, Shn-1 and Shn-2 have recently been

demonstrated to play a role in BMP-dependent transcriptional regulation of mesodermal patterning and adipogenesis, respectively (Yao et al., 2006; Jin et al., 2006). In this study, we have addressed the molecular mechanisms of SMA-9 function in BMP-related signaling in *C. elegans.*

C. elegans DBL-1 (Dpp and BMP-like) is a member of the TGF-β superfamily closely related to *Drosophila* Dpp and vertebrate BMP-2/BMP-4 (Suzuki et al., 1999). The DBL-1 signaling pathway regulates body size and patterning of male-specific copulatory structures (Savage-Dunn, 2005). DBL-1 pathway loss of function (*lf*) mutants have small body size in both hermaphrodites and males (Sma phenotype). In addition, the mutant male has abnormal male tail sensory ray fusions and crumpled spicules (Mab phenotype) (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 two sets of nine sensory rays distinguished by their specific positions and shapes. DBL-1 pathway (*lf*) mutant males have frequent ray fusions of rays 4-5, 6-7, and 8-9. Furthermore, in these mutants the pair of copulatory spicules necessary for mating are crumpled with 100% penetrance. When *dbl-1* is overexpressed in a wild-type background, the animal has a long body size (Lon phenotype) and the male contains ray 3-4 fusions (Suzuki et al., 1999). *sma-9* functions as a downstream component of the DBL-1 pathway. Like other DBL-1 pathway components, *sma-9*(*lf*) mutants have small body size and abnormal male tails. Phenotypic analysis shows that *sma-9* regulates body size development specifically in early larval stages and regulates morphogenesis of only sensory rays 8 and 9 (Liang et al., 2003). In addition to these roles in body size and male

tail patterning, *sma-9* has recently been reported to regulate mesodermal cell lineages, a role in which *sma-9* antagonizes DBL-1 pathway function (Foehr et al., 2006).

In *Drosophila*, early reports attributed to Shn roles as both a transcriptional activator and a transcriptional repressor (Dai et al., 2000; Marty et al., 2000; Torres-Vazquez et al., 2000). Some studies have suggested that the principle role of Shn is to repress expression of *brinker (brk)*, a transcriptional repressor of Dpp target genes (Marty et al., 2000; Muller et al., 2003). However, *brk* is not a conserved gene, and to date, no *brk* homolog has been found in *C. elegans* or other genomes. More recently, Pyrowolakis et al. reported that transcriptional repression of *brk* is dependent on a short *cis*-acting silencer element (SE) (Pyrowolakis et al., 2004). The SE has high affinity for the Smads Mad and Medea, which subsequently recruit 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, whether Shn contributes to the transcriptional repression of these genes has in most cases not been tested. In an elegant extension to this work, Yao et al. (2006) have reported that vertebrate Shn1 regulates the *Xenopus* BMP target gene *Xvent2* via a sequence homologous to the SE. In this system, Shn functions as a transcriptional activator rather than as a repressor. By exchanging *Drosophila* and *Xenopus* Shn genes in these experiments, they demonstrate that transcriptional activation or repression may be context-dependent rather than an intrinsic property of the protein.

None of these previous studies have determined whether transcriptional repressor or activator activities are sufficient to mediate Shn functions throughout development, rather than in the context of selected target genes, nor have studies addressed the intrinsic

transcriptional activities of Shn proteins. To address the nature of the SMA-9 transcriptional activity, we have used transgenic and DNA microarray approaches. These approaches are facilitated in *C. elegans* relative to other systems, such as *Drosophila* and vertebrates, because DBL-1 pathway mutants are viable. We have created SMA-9 fusions to known transcriptional activator and repressor domains. The transcriptional repressor fusion partially rescued *sma-9* mutant phenotypes in all tissues tested, and caused a ligand overexpression phenotype in the male tail. The transcriptional activator fusion acted as a dominant negative in body size, but partially rescued the male tail phenotypes. Additionally, we have shown that regions of SMA-9 have intrinsic transcriptional repressor and activator activities. Finally, we have used DNA microarray analysis to identify target genes of *dbl-1* and *sma-9* involved in early larval development, when *sma-9* is active in regulating body growth. This analysis expands substantially on a previous smaller-scale screen of transcriptional target genes for this pathway using arrayed cDNAs (Mochii et al., 1999). In our microarray positives, there are more genes transcriptionally repressed than activated. We identified conserved and novel transcriptional target genes, including homologs of Runx and bZIP transcription factors. Notably, we demonstrate that T27F2.4, a *C. elegans* homolog of the mouse gene fosB, is a downstream negative regulator of DBL-1 pathway-regulated development. Taken together, these experiments suggest that *sma-9* provides transcriptional repressor activity to regulate body size, while contributing both repressor and activator activities for the regulation of male tail patterning.

Materials and methods

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 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

En and VP-16 fusions: 4.0 kb *sma-9* promoter region (Liang et al., 2003) was cloned in pBluescript SK+ at Not I and Sac II sites (pCS301); and then 3.6 kb *sma-9* C-terminal genomic sequence, 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.

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. These constructs were cotransfected with pLGΔ312S and pJK1621 (gifts from S. Kuchin) into yeast strain MCY 829. Yeast transcription activity was performed as described (Kuchin and Carlson, 2003). Heat-shock vector pPD49.83 was a kind gift from Dr. Liu. *yk1285a11* and *yk1237d01* was cloned into pPD49.83 at Nhe I and Kpn I.

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

All primer sequences are available on request.

Transgenic animals and RNAi

Microinjection was done by standard procedures using *rol-6* as a marker (Mello et al., 1991). Arrays were integrated 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;* CS491- *him-5(e1490);qcIs26;* carrying pCS304 (*VP-16-sma-9)* are: CS488- *sma-9(wk55);him-5(e1490);qcIs33* 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. Transgenic animals carrying *yk1285a11* and *yk1237d01* are CS486 and CS482, respectively. dsRNA was made from pCS359 as described (Liang et al., 2003). RNAi feeding method using pCS308 and pCS360 was performed as described (Kamath and Ahringer, 2003). For the growth curve, RNAi hypersensitive *rrf-3* mutants were fed T27F2.4 dsRNA.

Heat-shock experiments.

Gravid hermaphrodites were washed off plates with M9 buffer and collected in an Eppendorf tube. Hermaphrodites were destroyed with sodium hypochlorite solution, leaving eggs that were collected and inoculated onto OP50-seeded plates. For male tail ray examination, 72 hours after egg collection, collect and transfer the worms to a siliconized Eppendorf tube containing 100ul of M9 buffer (Brenner, 1974), which was placed in a circulating waterbath under the heat-shock conditions specified. After heat-shock, worms were recovered, placed at 25°C on OP50-seeded plates and allowed to develop to adulthood and then scored for ray patterning.

DNA microarray analysis

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 until worms developed to L2 (staged by examining VPC and gonadal lineages using Nomarski optics); and harvested. Total RNA was extracted by standard methods using Trizol. Gene Chip *C. elegans* Genome Arrays 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 Affymetrix Company) and the Macro program (developed by GCL Memorial Sloan-Kettering Cancer Center). Microarray data have been submitted to ArrayExpress (accession number e-mexp-687).

Genes were selected that showed at least 2-fold changes (microarray difference P< 0.001) of expression level in pairwise comparison with wild type. To determine the accuracy of the microarray results, we performed quantitative real-time RT-PCR on 10 selected genes (Supplementary Table 1). For five of the ten genes, the RT-PCR results confirmed the microarray results for both genetic backgrounds, an additional three genes were confirmed in one of the two genetic backgrounds, and two genes were found to be false positives. We then performed a Student's t-test analysis on our microarray data. Elimination of genes with t-test P values >0.05 effectively removes one of the two false positives, *msp-152*, and inconsistent results for *rnt-1*. This stringency of statistical cutoff may also eliminate some true positives, such as T27F2.4 in the *dbl-1* background and *fat-7* in both backgrounds, but leaves us with greater confidence in the remaining group. Therefore, we report here genes that meet the criteria of fold change ≥2-fold and t-test P value ≤0.05 (Supplementary Tables 2-4).

RT-PCR

Total RNA was isolated by Trizol (GibcoBRL) from wild type, *sma-9(wk55),* the transformed strains CS332 and CS339. RT-PCR was performed using SUPERSCRIPT™ One-Step RT-PCR (Invitrogen). Real time PCR was performed on LightCycler 2.0 version (Roche). Primers were designed by LightCycler Probe Design Software 2.0. Data Analysis software used was LightCycler software 4.0 version. N2, *sma-9(wk55),* and *dbl-1 (wk70)* animals were synchronized to L2 as described above. Total RNA was extracted from these animals by Trizol as above. SuperScript™ III platinum two-step qRT-PCR kit with SYBR Green was used to perform RT-PCR. Totally 10 genes were examined both

in *sma-9* and *dbl-1* background. Gene *act-1* was used as standard control. The intron primer of *act-1* gene was used to test for the presence of genomic DNA. The results showed that total RNA was pure without contamination of genomic DNA.

Results

SMA-9 transcriptional repressor fusion rescues body size

To investigate whether *sma-9* acts as a transcriptional activator or repressor in the DBL-1 pathway, two artificial *sma-9* constructs were created, *VP-16-sma-9* and *en-sma-9*. *sma-9* genomic sequences encoding the zinc finger domains that are putative DNA binding domains were fused with either VP-16 or Engrailed (En) transcription activity domain, and expression driven by previously characterized *sma-9* promoter sequences. Since VP-16 and En have been shown to be a strong transcriptional activator and repressor, respectively (Sze et al., 1997; Tolkunova et al., 1998), 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 animals. Both the Sma and Mab phenotypes of the transformed animals were investigated. We used RT-PCR with transgene-specific primers to confirm expression of the fusion constructs (Fig. 2).

The *sma-9*(*lf*) mutants with *en-sma-9* constructs show significant rescue of the body size defect (Fig. 2A). In contrast, *VP-16-sma-9* transformed into either *sma-9*(*lf*) or wild-type animals results in a significant reduction in body size (Fig. 2A). These results suggest that for body size regulation, *en-sma-9* can substitute for functional *sma-9*, whereas *VP-16-sma-9* acts as a dominant negative.

Both SMA-9 transcriptional repressor and activator fusions partially rescue male tail defects

In *sma-9*(*lf*) mutants with *en-sma-9*, the male tail ray 8 and 9 fusions were improved from 53% frequency to 29%, and crumpled spicules were partially rescued from 46% frequency to 8% (Table 1). Meanwhile a *dbl-1* overexpression phenotype, ray 3 and 4 fusion, was observed (Fig. 2B). Unlike in body size regulation, *VP-16-sma-9* expression resulted in a partial rescue of *sma-9* male tail defects, reducing fusions from 53% to 36% (Table 1). Thus, in the male tail, repressor fusions and activator fusions can both partially substitute for endogenous SMA-9. We therefore hypothesize that SMA-9 protein products may provide both transcriptional repressor and activator functions, and that these functions may be differentially required in different tissues.

Transcriptional activity of SMA-9 protein fragments

To test whether SMA-9 protein products have intrinsic transcriptional activity, we performed a transcription activity assay in yeast (Table 2). SMA-9 fragments were chosen that contained possible transcriptional activity domains but not zinc finger motifs, to avoid intrinsic DNA binding activity. The cDNA clone *yk1285a11* contains the 5' Gln-rich domain, pCS272 contains 5' sequences but excludes the Gln-rich domain, and ARD II contains the second acidic residue-rich domain in predicted exons 14 and 15. The vector pLGΔ312S constitutively expresses the *lacZ* gene whose expression level can be measured by β-Galactosidase (β-Gal) activity. Vectors pJK1621 is a derivative that contains four LexA operators upstream of the UAS. The tested SMA-9 fragments were fused with LexA DNA binding domain in yeast vector pSH 2-1. After cotransformation,

the tested protein fragment should influence *lacZ* expression via interaction between LexA DNA binding domain and its operator. In the yeast assay, *yk1285a11* did not demonstrate transcriptional activity (data not shown). pCS272 decreased activity 2.3 fold (Table 2), demonstrating that this N-terminal fragment has transcriptional repressor activity. In contrast, the ARD II fragment causes a 1.2 fold increase of β-Gal expression (Table 2). This indicates that ARD II contains weak transcriptional activator activity. Interestingly, the cDNA clone *yk1237d01* contains an SL2 splice leader and polyA tail, representing a short but complete 3' *sma-9* transcript. Since this form excludes N-terminal sequences but contains ARD II and the putative DNA binding domain, it may have transcriptional activator activity in vivo (Fig. 1A).

To test the activities of these SMA-9 regions in vivo, we expressed the 5' and 3' cDNAs *yk1285a11* and *yk1237d01* in *sma-9(wk55)* using a heat-shock promoter. After heat shock, the 3' cDNA *yk1237d01*, but not the 5' clone *yk1285a11*, was able to provide rescuing activity in the male tail (Table 1). In fact, the degree to which *yk1237d01* rescues the frequency of ray 8-9 fusions is strikingly similar to the degree of rescue by the VP-16-SMA-9 fusion. These results are consistent with the hypothesis that this cDNA encodes a shorter SMA-9 isoform that specifically provides transcriptional activator activity in vivo.

More genes were repressed than were activated in DNA microarray expression profiles

To identify potential *dbl-1* and *sma-9* transcriptional target genes regulating body size, we determined patterns of gene expression using a DNA microarray. Since *sma-9* functions in early larval stages in body size regulation (Liang et al., 2003), L2 animals

were used to perform DNA microarray analysis. In this analysis, we would be unlikely to find target genes specifically involved in male tail development, since our populations did not include a significant proportion of males. Because the animals used are loss of function mutants, genes upregulated in the mutant backgrounds are repressed by the specific gene activity in the wild-type background, while genes downregulated in the mutant backgrounds are activated by the specific genes. The genes identified include both direct and indirect target genes.

We identified genes with ≥2-fold change in expression levels by microarray analysis (Materials and methods). Microarray positives were further culled by a Student's t-test, $p \leq 0.05$. Selected microarray positives were analyzed by quantitative RT-PCR (Supplementary Table 1) to confirm the validity of our statistical criteria. Based on RT-PCR results, these stringent criteria likely underestimate the total pool of target genes and, in particular, the overlap between *dbl-1* and *sma-9* regulated targets. The remaining pool, however, provide a core of target genes in which we have a high degree of confidence. By these criteria, a total of 242 genes have altered expression in *dbl-1* mutants and 146 genes in *sma-9* mutants, of which 31 genes are in common between both backgrounds (Fig. 3). Since *sma-9* regulates fewer genes than *dbl-1*, these data are consistent with our previous hypothesis that *sma-9* regulates a subset of DBL-1 pathway target genes. 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. A previously reported smaller-scale study of the DBL-1 pathway using arrayed cDNAs identified 24 genes commonly regulated by *dbl-1* and *sma-2* (Mochii et al, 1999). Four of these 24 genes were also identified in our screen (C08E3.13, F35C5.9, K01A2.3, K05G3.3; Supplementary Tables

2-4). Differences in the two sets of genes may be due in part to stage-specific differences in expression.

Among the 146 genes directly or indirectly regulated by *sma-9,* 107 genes are repressed and 39 genes are activated. These results are consistent with our hypothesis that *sma-9* functions as a transcriptional repressor in body size regulation. We expected that DBL-1 activity would cause the repression of some target genes, possibly mediated by SMA-9, and the activation of a greater number of target numbers. Surprisingly, among the 242 genes regulated by *dbl-1,* 159 genes are repressed but only 83 genes are activated. This suggests that in early larval stages in hermaphrodites, the DBL-1 pathway primarily represses target gene expression. We tested the hypothesis that the DBL-1 pathway represses target gene transcription by stimulation of *sma-9* expression, which then acts as a transcriptional repressor. Different *sma-9* transcriptional reporter genes (Liang et al., 2003) were introduced into *dbl-1, sma-6, sma-2, sma-3,* and *sma-4*. However, none of them had a significant expression change compared with wild type (data not shown). Also, *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 primarily as a transcriptional repressor in body size regulation.

Target genes involved in body size regulation

Five of our identified target genes have a demonstrated body size phenotype. One gene negatively regulates body size: the fos-related basic-leucine zipper transcription

factor T27F2.4, for which we have identified a Lon phenotype in RNAi experiments (see below). Another predicted transcription factor gene, *rnt-1,* regulates body size and was upregulated in *sma-9* mutants. *rnt-1* encodes the only homolog of mammalian *AML/RUNX* genes in *C. elegans* (Nam et al., 2002). Partial loss of function *rnt-1* mutants are small, and also their male tails have a broad range of defects in rays 1 to 9 but do not have a defect in the spicules (Yi et al., 2004; Nimmo et al., 2005). Both of these defects may be due in part to disruptions in cell divisions of the seam cells (Nimmo et al., 2005). Mammalian Runx genes, also known as acute myeloid leukemia (AML), encode the DNA-binding α chain of the heterodimer core-binding factor (CBF) complex. Runx1 is a frequent target of chromosome translocation in human leukemia especially childhood acute myeloid leukemia (Mitani, 2004). Runx2 has been reported to interact with Smads and regulate TGF-β type I receptor expression (Selvamurugan et al., 2004; Ji et al., 1998). In addition, TGF-β activates Runx genes at both transcriptional and post-transcriptional levels (Jin et al., 2004; Lee et al., 2002). Our RT-PCR results confirmed upregulation of *rnt-1* in *dbl-1* and *sma-9* mutant backgrounds (Supplementary Table 1). Thus, we propose that *rnt-1* may be a downstream target gene of the DBL-1 pathway. Since 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 (Yi et al., 2004) and *rnt-1* mutants have additional defects (Nimmo et al., 2005), we suggest that additional inputs other than the DBL-1 pathway also impinge on *rnt-1* function.

Our data showed two hedgehog-related genes and two patched-related genes: *wrt-4, wrt-9*, *ptc-3,* and *ptr-22,* regulated by the DBL-1 pathway (Table 3). Among these four genes, three have been shown to regulate body size (Zugasti et al., 2005). The Hedgehog

(Hh) family and its receptors (patched and smoothened) play an essential role in *Drosophila* and vertebrate development (Tabata and Takei, 2004). The genetic interaction between Hh and BMPs has been extensively studied in other organisms (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 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*). Of particular interest among these genes is *wrt-4*, whose change in expression levels was significant in both *dbl-1* and *sma-9* mutants and reproducible by RT-PCR (Supplementary Table 1). Interestingly, *wrt-4* and *ptc-3* were also listed in TGF-β dauer pathway regulated genes (Liu et al., 2004), suggesting crosstalk between these pathways.

Expression of the bZIP gene T27F2.4 is upregulated in DBL-1 pathway mutants

In both the microarray and RT-PCR analyses, the uncharacterized bZIP gene T27F2.4 was upregulated in *dbl-1* and *sma-9* mutant backgrounds, although the t-test showed marginal significance for the microarray expression change in *dbl-1* ($P = 0.1024$). To characterize this gene's regulation further in vivo we created a reporter construct. Significantly, T27F2.4p::GFP has a stronger expression in *sma-9, dbl-1,* and *sma-6*(*lf*) mutants compared with wild type (Fig. 4A-E). This result supports the interpretation that T27F2.4 is an actual target of the DBL-1 pathway. Both in wild type and DBL-1 pathway

mutants, T27F2.4p::GFP expression in neuronal cells was first observed at the L1 larval stage and then slowly disappeared after L4 (Fig. 4B). In wild type, a very weak expression in intestine was observed only in adults. However, in all DBL-1 pathway mutants a much stronger expression in intestine occurred after L4 stages (Fig. 4A, C, D). Interestingly, a weak expression in intestine was observed only in *sma-9* mutant background during early larval stages, possibly correlating with the highly significant change in expression in this mutant background at the L2 stage when examined by microarray analysis. The intestine is a common site of expression for DBL-1 pathway signaling components (Savage-Dunn, 2005) and identified target genes (Mochii et al., 1999), though it is not the focus of body size regulation (Inoue and Thomas, 2000; Yoshida et al., 2001; Wang et al., 2002). Based on only a transcriptional reporter construct, however, we cannot exclude the possibility that T27F2.4 is also expressed elsewhere, such as in the hypodermis.

To further investigate the function of T27F2.4 in the DBL-1 pathway, dsRNAi was performed (Fig. 4). After microinjection of T27F2.4 dsRNA in *him-5*, body size was slightly longer than control (Fig. 4G); and male tails contained a ray 3-4 fusion similar to that displayed in ligand overexpression males (data not shown). To further quantify the body size phenotype, we used RNAi by feeding. In the wild-type background, these animals were the same as controls (data not shown). In the *rrf-3* RNAi hypersensitive background (Simmer et al., 2002), T27F2.4 knockdown resulted in a long body size in adults and larvae after about 24 hours of development (Fig.4F).

T27F2.4 encodes a basic-leucine zipper transcription factor homologous to the fos homologs human $p21^{SNFT}$ and mouse fosB. Fos proteins are dimerization partners with

Jun proteins that bind AP-1 binding sites. JunB is known to be a Smad-binding partner, and both fos and jun homologs are transcriptionally regulated by TGF-β signaling (Jonk et al., 1998; Selvamurugan et al., 2004). In addition, the bZIP protein ATF3 is transcriptionally regulated by Smads and subsequently act as a Smad transcriptional co-factor (Kang et al., 2003). In *C. elegans*, four bZIP genes with similarity to vertebrate fos are identifiable: *atf-1*, *ces-1*, *fos-1*, and T27F2.4, although only *fos-1* may be a true fos ortholog (Sherwood et al., 2005). Among these, the mutant phenotype of T27F2.4 had not previously been characterized. Our RNAi and expression analyses indicate that T27F2.4 is repressed by the DBL-1 pathway and functions as a negative regulator of DBL-1 regulated body size.

Discussion

SMA-9 transcriptional repressor and activator activities function differentially in DBL-1 pathway outputs

sma-9 encodes zinc finger transcription factors homologous to *Drosophila* Schnurri. We have characterized the transcriptional activities of SMA-9 and conclude that it acts primarily as a transcriptional repressor in body size regulation, but as both a transcriptional repressor and activator in male tail patterning. First, we have shown that the En transcriptional repressor fused to SMA-9 DNA binding domains can substitute for functional SMA-9 protein in vivo (Table 1). 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. Second, we have shown that VP-16-SMA-9 is dominant negative in body size regulation but partially rescues male tail

defects (Fig. 2; Table 1). Third, we showed that different *sma-9* fragments have different transcriptional activities in vitro (Table 2). Given that alternative splicing exists in *sma-9*, isoforms containing the N-terminus may act as transcriptional repressors, while some isoforms, for example that encoded by *yk1237d01,* could function as transcriptional activators in the pathway. Consistent with this hypothesis, *yk1237d01* expression can rescue the sensory ray defect of *sma-9(wk55)* to the same extent as the activator fusion VP-16-SMA-9. The fourth line of evidence comes from our microarray analysis. Among genes directly or indirectly regulated by *sma-9,* more genes are repressed than are activated, consistent with *sma-9* functioning primarily as a transcriptional repressor in the pathway in early larval stages. Work on *Drosophila* and *Xenopus* Schnurri homologs has demonstrated that transcriptional activation or repression at specific target genes may be context-dependent (Yao et al., 2006). Our results do not preclude a context-dependent mechanism, but suggest that intrinsic properties of the protein products also contribute to the determination of transcriptional activity.

Our results suggest that transcriptional repression is critical to the response to DBL-1. Specifically, among *dbl-1* regulated genes, 66% of genes were repressed, suggesting that in early larval stages the DBL-1 pathway primarily represses target genes, whether they are SMA-9-dependent or -independent. At a molecular level, this repression may be mediated by Smad complexes directly, or by recruitment of transcriptional repressors such as SMA-9, or by the transcriptional activation of transcriptional repressor gene(s) that then repress target genes. In our findings, *sma-9* transcription was not regulated by the DBL-1 pathway itself, but SMA-9 might be recruited by Smad proteins in order to repress gene transcription. Our work underscores the importance of

investigating both negative and positive regulation of target genes in response to signaling pathways.

Genes that regulate body size

The molecular mechanisms by which the DBL-1 pathway regulates growth are incompletely understood. Two hypotheses in the literature are that the DBL-1 pathway regulates body size through controlling collagens and through regulating chromosomal ploidy due to endoreduplication in hypodermal cells (Nystrom et al., 2002; Flemming et al., 2000; Morita et al., 2002). In our microarray profile, we identified both cell cycle control genes and collagen genes that might regulate endoreduplication (see below). We also identified five target genes whose loss of function results in demonstrated body size phenotypes: *ptc-3*, *ptr-22*, *wrt-4*, T27F2.4 (fos-related), and *rnt-1*. The latter two are transcription factors that could mediate indirect modulation of gene expression downstream of DBL-1 signaling. Furthermore, these transcription factors have vertebrate homologs, fos and Runx, which interact directly with Smad proteins and are transcriptional targets of TGF-β-related signaling (Jonk et al., 1998; Selvamurugan et al., 2004). The loss-of-function phenotype of T27F2.4 is long body size, suggesting that the negative regulation of this gene may mediate DBL-1 signaling outputs. On the other hand, the loss-of-function phenotype of *rnt-1* (small body size) and its genetic interactions are not consistent with a simple model in which *rnt-1* regulates body size solely as a downstream target of *dbl-1*. Other conserved interactors that we identified were genes of the Hedgehog family. Genetic interactions between BMP pathways and Hedgehog signaling have been extensively studied in other organisms (Tabata and Takei, 2004),

however little is known in *C. elegans*. Finally, some novel targets of DBL-1 signaling were also identified. These include a family of nine genes containing a domain of unknown function, DUF19 (Supplementary Table 4).

Cell cycle control genes

mcm-7 and cell cycle control genes were downregulated in *dbl-1* but not *sma-9* mutants (Table 3). *mcm-7* is homologous to minichromosome maintenance (MCM) proteins that play a critical role in S phase genome stability. MCM are DNA replication licensing factors in eukaryotic 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).

We and others have shown that hyp7, the large hypodermal syncytium, is a critical tissue for body size development, since 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). 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., 2000). However, several lines of evidence showed that the DNA content in polyploid hypodermal nuclei due to endoreduplication was reduced in the pathway mutants (Flemming et al., 2000; Morita et al., 2002). Perhaps downregulation of *mcm-7* blocks endoreduplication in hypodermal cells. Alternatively, the low expression level of *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.

Collagens

Collagen genes are a large class of *dbl-1*-regulated, but *sma-9*-independent, target genes potentially involved in body size regulation (Fig. 3; Table 3). 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 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). In our data *rol-6* is downregulated in *dbl-1* mutants, which is consistent with these findings.

Genes related to innate immunity

In this study, we did not expect to identify target genes involved in male tail patterning, since we analyzed expression patterns in hermaphrodites during early larval stages. In addition to genes involved in body size regulation, however, we likely identified genes with functions in other outputs of DBL-1 signaling, such as innate

immunity, aging, and germline regulation (Supplementary Tables 2-4). 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 (Mallo et al., 2002; Nicholas and Hodgkin, 2004) (Fig. 3; Table 3). Similar genes were identified in a previous smaller-scale study of this pathway using arrayed cDNAs (Mochii et al., 1999).

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* (Supplementary Tables 2-4) including the genes F35C5.9, B0218.8, and T07D10.4 (Table 3) identified from previous studies of this pathway or of dauer pathways (Murphy et al., 2003; Mochii et al., 1999). Compared with previous reports on innate immune response (Mallo et al., 2002; Couillault et al., 2004), *nlp-29* and *lys-10* expression changed in the opposite direction. However, *nlp-29* was induced specifically upon bacterial *S. marcescens* or fungal *D. coniospora* infection of adult animals (Couillault et al., 2004), which represents a pathogen specific response in *C. elegans,* so expression may depend on culture conditions. 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)* mutants, together with *mrp-1* or *pgp-3,* were hypersensitive to heavy metals and bacteria,

indicating that *pgp-1* plays a role in a defense response. The function of *pgp-6* is still not characterized. In our analysis, all PGP genes were downregulated in the *dbl-1(lf)* mutant, which might contribute to the high susceptibility to bacterial infection of mutants.

Aging genes

We see altered transcription of genes identified in previous studies on dauer and aging (Murphy et al., 2003; Liu et al., 2004; Halaschek-Wiener et al., 2005; Table 3). In our array, genes extending lifespan were downregulated such as *fat-6;* genes shortening lifespan were upregulated such as *vit-5* and *vit-6*. 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 loss of function of the DBL-1 pathway leads to a shorter lifespan in a diet of OP50.

Germline genes

In *Drosophila,* genetic evidence has been provided that BMP homologs *gbb* and *dpp* are essential for maintaining germline stem cells (GSC) in the ovary and testis (Kawase et al., 2004; Song et al., 2004). However, in *C. elegans,* no functional relationship between BMP signaling and germline development has been previously proposed. Our microarray data includes many genes affecting germline development, including key genes required for meiosis, such as *gld-1* (Table 3). These meiosis-promoting genes are repressed by *dbl-1* and *sma-9*, as would be expected if *dbl-1* functioned to maintain stem cell identity. These results point to an evolutionarily

conserved role for DBL-1 pathway signaling in the germline that had not previously been recognized.

Acknowledgements

We thank Dr. Gary Struhl for the VP-16 and En vectors; Dr. Sergei Kuchin for the yeast vectors and strains; Dr. Agnes Viale (Genomics Core Facility, MSKCC) for microarray analysis; Dr. Richard Kolesnick and Dr. Xinzhu Deng for use of a cesium source for γ-irradiation; Dr. Nathalia Glickman Holtzman for comments on the manuscript. This research was supported by RPG 98-230-01-DDC and RSG-98-230-04-DDC from the American Cancer Society to C.S.-D. Some of the experiments were done with equipment from the Core Facilities for Imaging, Cellular and Molecular Biology at Queens College. Some *C. elegans* mutant strains were obtained from the Caenorhabditis Genetics Center, which is supported by the NIH National Center for Research Resources (NCRR). This work was carried out in partial fulfillment of the requirements for the Ph.D. degree from the Graduate Center of City University of New York (J.L.).

References

Aspock, G., Kagoshima, H., Niklaus, G., Burglin, T.R. 1999. Caenorhabditis elegans has scores of hedgehog-related genes: sequence and expression analysis. Genome Res. 9, 909-923.

Bailis, J.M., Forsburg, S.L. 2004. MCM proteins: DNA damage, mutagenesis and repair. Curr Opin Genet Dev. 14, 17-21.

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

Couillault, C., Pujol, N., Reboul, J., Sabatier, L., Guichou, J.F., Kohara, Y., Ewbank, J.J. 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, 488-494.

Dai, H., Hogan, C., Gopalakrishnan, B., Torres-Vazquez, J., Nguyen, M., Park, S., Raftery, L. A., Warrior, R. 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.

Derynck, R., Zhang, Y.E. 2003. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. Nature 425, 577-584.

Dupont, S., Zacchigna, L., Cordenonsi, M., Soligo, S., Adorno, M., Rugge, M., Piccolo, S. 2005. Germ-layer specification and control of cell growth by Ectodermin, a Smad4 ubiquitin ligase. Cell 121, 87-99.

Flemming, A.J., Shen, Z.Z., Cunha, A., Emmons, S.W., Leroi, A.M. 2000. Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. Proc Natl Acad Sci U S A 97, 5285-5290.

Foehr, M.L., Lindy, A.S., Fairbank, R.C., Amin, N.M., Xu, M., Yanowitz, J., Fire, Z.A., Liu, J. 2006. Development 133, 2887-2896.

Halaschek-Wiener, J., Khattra, J.S., McKay, S., Pouzyrev, A., Stott, J.M., Yang, G.S., Holt, R.A., Jones, S.J.M., Marra, M.A., Brooks-Wilson, A.R., Riddle, D.L. 2005. Analysis of long-lived *C. elegans daf-2* mutants using serial analysis of gene expression. Genome Research 15, 603-615.

He, W., Dorn, D.C., Erdjument-Bromage, H., Tempst, P. Moore, M.A., Massague, J. 2006. Hematopoiesis controlled by distinct TIF1gamma and Smad4 branches of the TGFbeta pathway. Cell 125, 929-941.

Hicar, M.D., Liu, Y., Allen, C.E., Wu, L.-C. 2001. Structure of the human zinc finger protein HIVEP3: molecular cloning, expression, exon-intron structure, and comparison with paralogous genes HIVEP1 and HIVEP2. Genomics 71, 89-100.

Higgins, C.F., Linton, K.J. 2004. The ATP switch model for ABC transporters. Nat Struct Mol Biol. 11, 918-926.

Hobert, O. 2002. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. Biotechniques 32, 728-730.

Inoue, T., Thomas, J.H. 2000. Targets of TGF-beta signaling in Caenorhabditis elegans dauer formation. Dev. Biol. 217, 192-204.

Ji, C., Casinghino, S., Chang, D.J., Chen, Y., Javed, A., Ito, Y., Hiebert, S.W., Lian, J.B., Stein, G.S., McCarthy, T.L., Centrella M. 1998. CBFa(AML/PEBP2)-related elements in the TGF-beta type I receptor promoter and expression with osteoblast differentiation. J Cell Biochem. 69, 353-363.

Ji, Y.J., Nam, S., Jin, Y.H., Cha, E.J., Lee, K.S., Choi, K.Y., Song, H.O., Lee, J., Bae, S.C., Ahnn, J. 2004. RNT-1, the C. elegans homologue of mammalian RUNX transcription factors, regulates body size and male tail development. Dev Biol. 274, 402-412.

Jin, W., Takagi, T., Kanesashi, S.N., Kurahashi, T., Nomura, T., Harada, J., Ishii, S. 2006. Schnurri-2 controls BMP-dependent adipogenesis via interaction with Smad proteins. Dev Cell 10, 461-471.

Jin, Y.H., Jeon, E.J., Li, Q.L., Lee, Y.H., Choi, J.K., Kim, W.J., Lee, K.Y., Bae, S.C. 2004. Transforming growth factor-beta stimulates p300-dependent RUNX3 acetylation, which inhibits ubiquitination-mediated degradation. J Biol Chem. 279, 29409-29417.

Jonk, L.J.C., Itoh, S., Heldin, C.H., ten Dijke, P., Kruijer, W. 1998. Identification and functional characterization of a Smad binding element (SBE) in the *JunB* promoter that acts as a Transforming Growth Factor-β, Activin, and Bone Morphogenetic Protein-inducible enhancer. J. Biol. Chem. 273, 21145-21152.

Kamath, R.S., Ahringer, J. 2003. Genome-wide RNAi screening in Caenorhabditis elegans. Methods 30, 313-321.

Kang, Y., Chen, C.R., Massague, J. 2003. A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. Mol Cell 11, 915-926.

Kawase, E., Wong, M.D., Ding, B.C., Xie, T. 2004. Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing bam transcription in the Drosophila testis. Development 131, 1365-1375.

Kramer, J.M., French, R.P., Park, E.C., Johnson, J.J. 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, 2081-2089.

Kuchin, S., Carlson, M. 2003. Analysis of transcriptional repression by Mig1 in Saccharomyces cerevisiae using a reporter assay. Methods Enzymol. 371, 602-614.

Kurz, C.L., Tan, M.W. 2004. Regulation of aging and innate immunity in C. elegans. Aging Cell 3, 185-193.

Lee, K.S., Hong, S.H., Bae, S.C. 2002. Both the Smad and p38 MAPK pathways play a crucial role in Runx2 expression following induction by transforming growth factor-beta and bone morphogenetic protein. Oncogene 21, 7156-7163.

Liang, J., Lints, R., Foehr, M.L., Tokarz, R., Yu, L., Emmons, S.W., 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, 6453-6464.

Liu, T., Zimmerman, K.K., Patterson, G.I. 2004. Regulation of signaling genes by TGFbeta during entry into dauer diapause in C. elegans. BMC Dev Biol. 4, 11.

Loris, R. 2002. Principles of structures of animal and plant lectins. Biochim Biophys Acta. 1572, 198-208.

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

Mallo, G.V., Kurz, C.L., Couillault, C., Pujol, N., Granjeaud, S., Kohara, Y., Ewbank, J.J. 2002. Inducible antibacterial defense system in C. elegans. Curr Biol. 12, 1209-1214.

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

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

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

Mitani, K. 2004. Molecular mechanisms of leukemogenesis by AML1/EVI-1. Oncogene 23, 4263-4269.

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, 15020-15025.

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

Muchardt, C., Seeler, J.S., Nirula, A., Shurland, D.L., Gaynor, R.B. 1992. Regulation of human immunodeficiency virus enhancer function by PRDII-BF1 and c-rel gene products. J. Virol., 66, 244-250.

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

Murphy, C.T., McCarroll, S.A., Bargmann, C.I., Fraser, A., Kamath, R.S., Ahringer, J., Li, H., Kenyon, C. 2003. Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature 424, 277-283.

Myllyharju, J., Kivirikko, K.I. 2004. Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet. 20, 33-43.

Nam, S., Jin, Y.H., Li, Q.L., Lee, K.Y., Jeong, G.B., Ito, Y., Lee, J., Bae, S.C. 2002. Expression pattern, regulation, and biological role of runt domain transcription factor, run, in Caenorhabditis elegans. Mol Cell Biol. 22, 547-554.

Nicholas, H.R., Hodgkin, J. 2004. Responses to infection and possible recognition strategies in the innate immune system of Caenorhabditis elegans. Mol Immunol. 41, 479-493.

Nimmo, R., Antebi, A., Woollard, A. 2005. mab-2 encodes RNT-1, a C. elegans Runx homologue essential for controlling cell proliferation in a stem cell-like developmental lineage. Development 132, 5043-5054.

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

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

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, 229-240.

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

Savage, C., Das, P., Finelli, A.L., Townsend, S.R., Sun, C.Y., Baird, S.E., 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, 790-794.

Savage-Dunn, C. 2005. TGF-β signaling (September 9, 2005), *WormBook*, ed. The *C. elegans* Research Community, doi/10.1895/wormbook.1.22.1, http://www.wormbook.org

Selvamurugan, N., Kwok, S., Partridge, N.C. 2004. Smad3 interacts with JunB and Cbfa1/Runx2 for transforming growth factor-beta1-stimulated collagenase-3 expression in human breast cancer cells. J Biol Chem. 279, 27764-27773.

Sherwood, D.R., Butler, J.A., Kramer, J.M, Sternberg, P.W. 2005. FOS-1 promotes basement-membrane removal during anchor-cell invasion in C. elegans. Cell 121, 951-962.

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

Silver, S.J., Rebay, I. 2005. Signaling circuitries in development: insights from the retinal determination gene network. Development 132, 3-13.

Simmer, F., Tinsterman, M., Parrish, S., Koushika, S.P., Nonet, M.L., Fire, A., Ahringer, J., Plasterk, R.H. 2002. Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Current Biology 12, 1317-1319.

Song, X., Wong, M.D., Kawase, E., Xi, R., Ding, B.C., McCarthy, J.J., 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, 1353-1364.

Sulston, J.E., Albertson, D.G., Thomson, J.N. 1980. The Caenorhabditis elegans male: postembryonic development of nongonadal structures. Dev Biol. 78, 542-576.

Suzuki, Y., Morris, G.A., Han, M., Wood, W.B. 2002. A Cuticle Collagen Encoded by the lon-3 Gene May Be a Target of TGF-beta Signaling in Determining Caenorhabditis elegans Body Shape. Genetics 162, 1631-1639.

Suzuki, Y., Yandell, M.D., Roy, P.J., Krishna, S., Savage-Dunn, C., Ross, R.M., Padgett, R.W., 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.

Sze, J.Y., 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, 1159-1168.

Tabata, T., Takei, Y. 2004. Morphogens, their identification and regulation. Development 131,703-712.

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

Tolkunova, E.N., Fujioka, M., Kobayashi, M., Deka, D., Jaynes, J.B. 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, 2804-2814.

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

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

Wu, L.C. 2002. ZAS: C2H2 zinc finger proteins involved in growth and development. Gene Expr. 10, 137-152.

Yao, L.-C., Blitz, I.L., Peiffer, D.A., Phin, S., Wang, Y., Ogata, S., Cho, K.W.Y., Arora, K., Warrior, R. 2006. Schnurri transcription factors from *Drosophila* and vertebrates can mediate Bmp signaling through a phylogenetically conserved mechanism. Development 133, 4025-4034.

Yoshida, S., Morita, K., Mochii, M., 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.

Zugasti, O., Rajan, J., Kuwabara, P.E. 2005. The function and expansion of the Patched- and Hedgehog-related homologs in C. elegans. Genome Res. 15, 1402-1410.

Figure legends

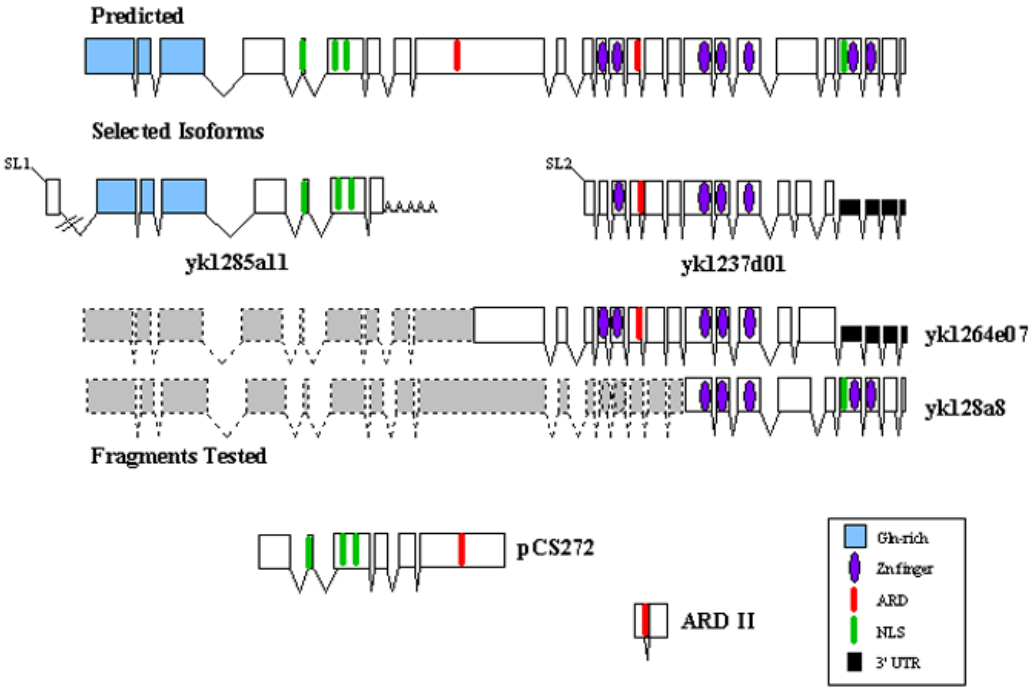


Fig. 1. Schematic of DNA constructs used in this study. *sma-9* cDNA fragments (pCS272 and ARDII) that were examined for transcriptional activity. The predicted *sma-9* genomic structure and structure of representative cDNA clones is shown above.

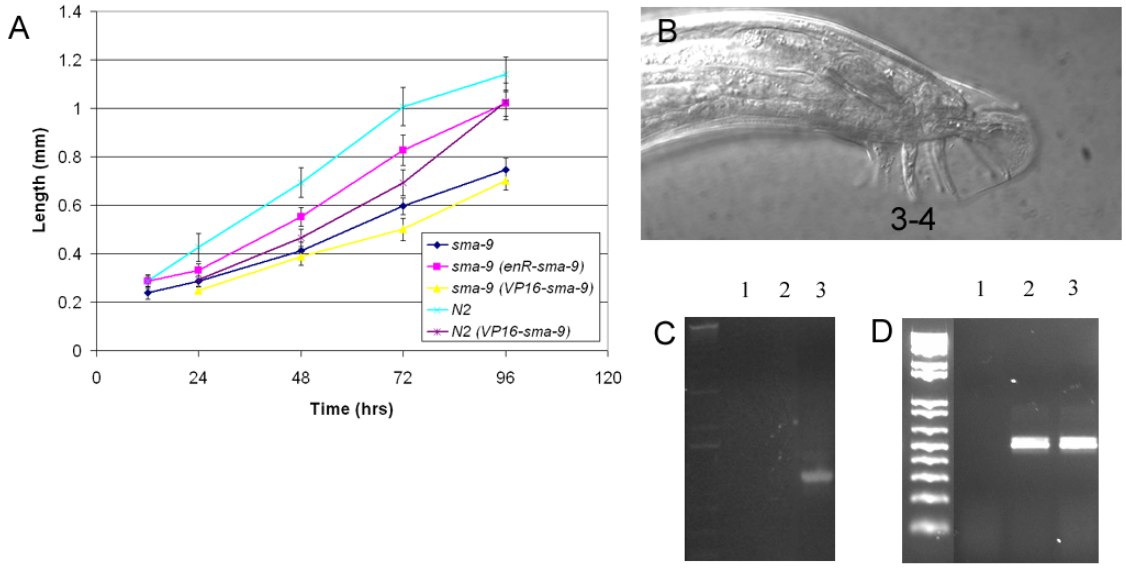


Fig. 2. *sma-9* functions as a transcriptional repressor in DBL-1 pathway regulated body size. (A) Growth curve of *sma-9(wk55)* animals transformed with *en-sma-9* (square), *VP-16-sma-9* (triangle), and control (diamond); and wild-type animals transformed with *VP-16-sma-9* (star) and control (cross). The *en-sma-9* transformed animals are larger than *sma-9*, indicating that it partially rescues the body size. Meanwhile, *VP-16-sma-9* transformed animals are smaller than control animals. (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. (C) Detection of *en-sma-9* mRNA. (D) Detection of *VP-16-sma-9* mRNA. The transcript from fusion constructs was detected by RT-PCR in transgenic strains *en-sma-9* (C-3), *sma-9;VP-16-sma-9* (D-2), and *VP-16-sma-9* (D-3), but not in untransformed wild type (C-1; D-1) backgrounds. *en-sma-9* expression was not detected in the *sma-9* mutant background (C-2), but rescue was observed in this strain, suggesting that the transgene was expressed at a low but sufficient level.

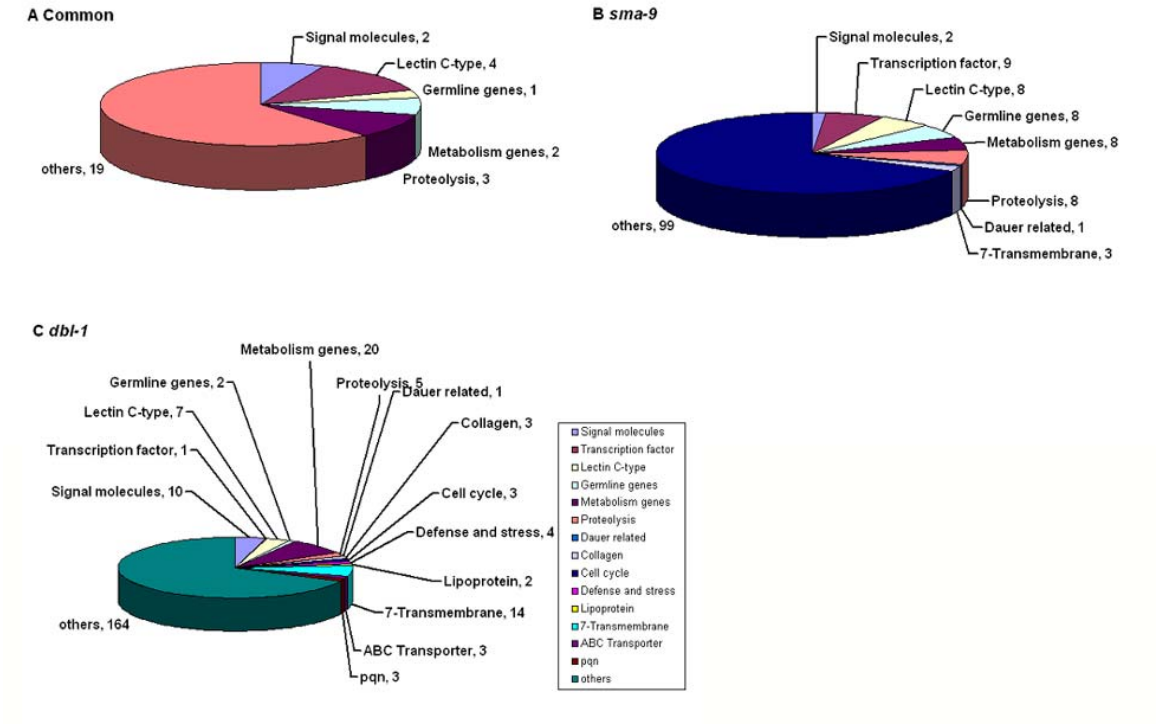


Fig. 3. Classes of genes identified from DNA microarray analysis regulated by *dbl-1* and *sma-9*. (A) Common genes regulated by *dbl-1* and *sma-9*. (B) Genes whose expression is altered in *sma-9* mutants. (C) Genes whose expression is altered in *dbl-1* mutants.

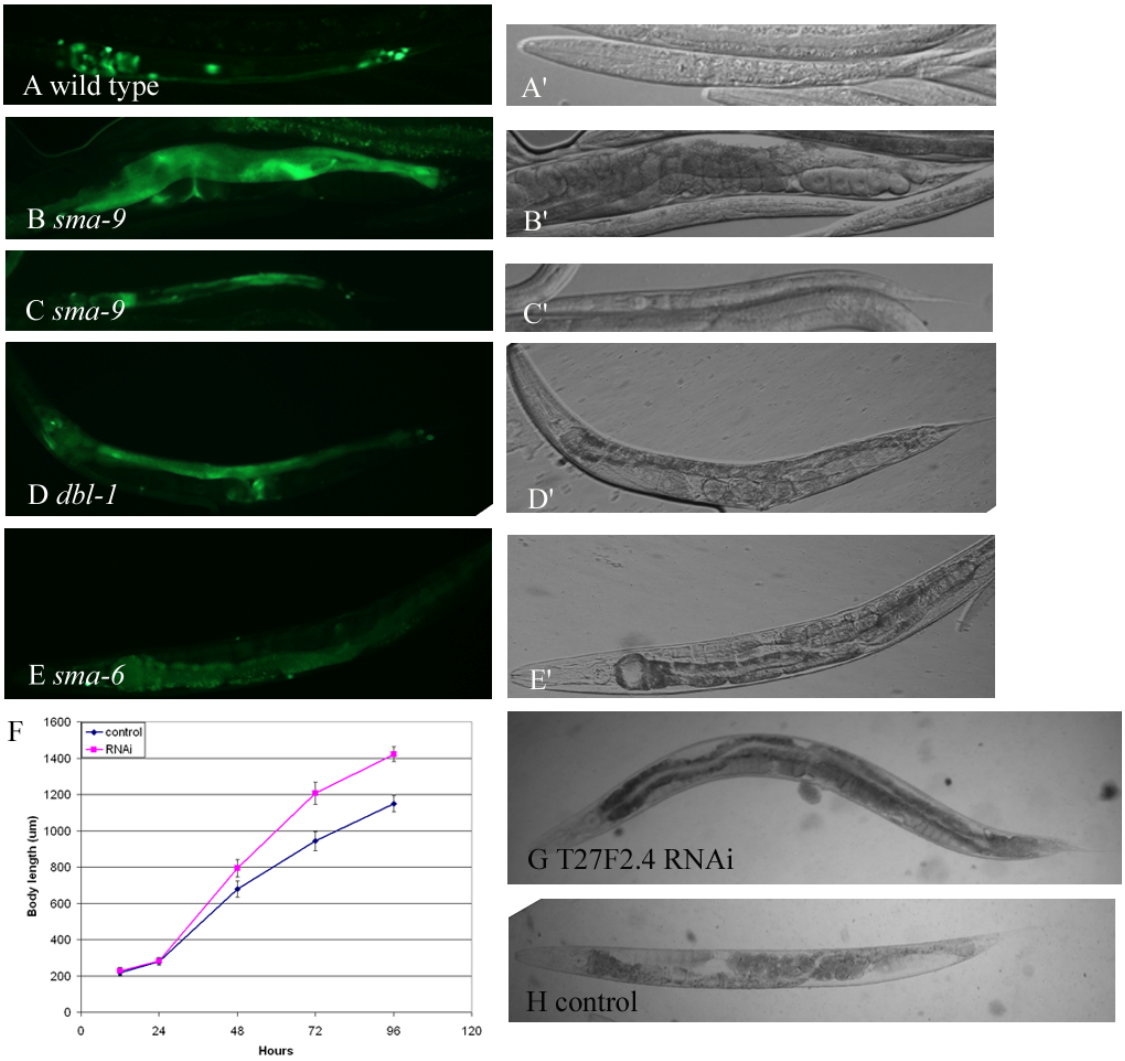


Fig. 4. T27F2.4 negatively regulates body size and male tail morphogenesis downstream of the DBL-1 pathway. (A-E) T27F2.4 expression level was significantly changed in DBL-1 pathway mutants. In wild type (A), neuronal expression were observed at larval stages but disappeared in adulthood. The same neuronal expression was displayed in DBL-1 pathway mutants (data not shown). T27F2.4 intestine expression was almost undetectable in wild type background. However, in *sma-9(wk55)* background a high intestine expression was observed in adult (B); while a low intestine expression was detected in larval stages (C). *dbl-1(wk70)* animal expressed a similar level of T27F2.4 in

intestine (D) as in *sma-9*. In *sma-6(wk7)* mutants, the intestine expression is lower compared with *dbl-1* and *sma-9,* however it is still significantly higher than in wild type (E). (A'-E') Nomarski images of animals in (A-E). (F-H) (F) Growth curve showing long body size phenotype of T27F2.4 RNAi animals. T27F2.4 dsRNAi adults (G) are longer than control (H), indicating that it might be a negative regulator of DBL-1 pathway regulated body size.

Table 1

Male tail phenotypes in *sma-9* transgenic animals

Strain	Ray 3-4 fusion	Ray 8-9 fusion	n (sides)	Crumpled spicules	n (tails)	Ectopically expressed product
wild type	0%	16%	74	0%	30	None
qcIs26	Occasionally Observed	11%	88	0%	52	En-SMA-9
qcIs33	0%	8%	74	0%	30	VP-16-SMA-9
sma-9	0%	53%	101	46%	50	None
sma-9;qcIs26	Occasionally Observed	29%**	87	8%**	36	En-SMA-9
sma-9;qcIs33	0%	36%*	50	ND	ND	VP-16-SMA-9
sma-9;qcEx118	0%	70%	20	ND	ND	*yk1285a11* (5')
sma-9;qcEx119	0%	34%**	85	ND	ND	*yk1237d01* (3')

All strains contain *him-5(e1490).* The *sma-9* allele used is *wk55*.

*Significant rescue: $p < 0.05$ compared to *sma-9* mutant background.

**Highly significant rescue: $p < 0.01$ compared to *sma-9* mutant background.

Table 2

SMA-9 fragment transcriptional activity

Fragment	β-Gal activity[1]		Fold Change[2]
	- LexA op	+ LexA op	
LexA	99	86	-1.15
pCS272-LexA	871.1 ± 61.1	378.9 ± 26.7	-2.3
ARD II-LexA	1512.7 ± 198.6	1826.3 ± 346.3	+1.2

[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.

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

Table 3

Target genes of the DBL-1 pathway identified by microarray analysis

Gene Name	Identification	Average fold change	Background	Reference[1]
Body size regulation				
T27F2.4	fos-related basic-leucine zipper (bZIP) transcription factor	4.33	*sma-9*	
rnt-1	homolog of mammalian *AML/RUNX* genes	3.31	*sma-9*	
ptc-3	patched related family	-2.38	*dbl-1*	Liu et al., 2004
ptr-22	patched related family	3.22	*dbl-1*	
wrt-4	warthog	-2.95	*dbl-1* and *sma-9*	Liu et al., 2004
Other hedgehog/patched				
wrt-9	warthog	-2.14	*dbl-1*	
Cell cycle				

C04C3.4	G2/Mitotic-specific cyclin A	2.75	*dbl-1*	
T07F10.5	G2/Mitotic-specific cyclin A	2.64	*dbl-1*	
mcm-7	DNA replication licensing factor	-2.07	*dbl-1*	
Collagens				
col-141	cuticular collagen (type IV and type XIII)	3.19	*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*	
Innate immunity				
nlp-29	neuropeptide-like protein	3.38	*sma-9*	Mallo et al., 2002
pgp-1	multidrug/pheromone exporter, ABC superfamily	-2.57	*dbl-1*	
B0218.8	C-type lectin	-2.79	*dbl-1* and *sma-9*	Murphy et al., 2003

pgp-6	multidrug/pheromone exporter, ABC superfamily	-2.89	*dbl-1*	
T07D10.4	C-type lectin	-2.95	*dbl-1*	Murphy et al., 2003
F35C5.9	C-type lectin	-3.85	*dbl-1* and *sma-9*	Mochii et al., 1999
lys-10	lysozyme	-4.27	*sma-9*	
Aging genes				
vit-5	yolk protein	15.51	*dbl-1*	Murphy et al., 2003; Halaschek-Wiener et al., 2005
vit-6	yolk protein	5.29	*dbl-1*	Halaschek-Wiener et al., 2005
fat-6	acyl-CoA desaturase	-3.25	*dbl-1*	Murphy et al., 2003; Halaschek-Wiener et

al., 2005

Germline

gld-3	defective in germ line development	19.59	*dbl-1*
B0280.5	a direct target of GLD-1	3.70	*sma-9*
gld-1	defective in germ line development	2.20	*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*
fog-3	feminization of germline	2.00	*sma-9*
glh-2	germ-line helicase	2.00	*sma-9*
cav-1	caveolin protein family	-2.98	*dbl-1* and *sma-9*

[1]other published expression studies

References:

Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 72:291-336

Blumenthal T. 2005. Trans-splicing and operons. *WormBook*:1-9

Blumenthal T, Evans D, Link CD, Guffanti A, Lawson D, et al. 2002. A global analysis of Caenorhabditis elegans operons. *Nature* 417:851-4

Bosher JM, Dufourcq P, Sookhareea S, Labouesse M. 1999. RNA interference can target pre-mRNA: consequences for gene expression in a Caenorhabditis elegans operon. *Genetics* 153:1245-56

Brenner S. 1974. The genetics of Caenorhabditis elegans. *Genetics* 77:71-94

Chen X, Rubock MJ, Whitman M. 1996. A transcriptional partner for MAD proteins in TGF-beta signalling. *Nature* 383:691-6

Chen X, Weisberg E, Fridmacher V, Watanabe M, Naco G, Whitman M. 1997. Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature* 389:85-9

Chen YG, Hata A, Lo RS, Wotton D, Shi Y, et al. 1998. Determinants of specificity in TGF-beta signal transduction. *Genes Dev* 12:2144-52

Choi J, Newman AP. 2006. A two-promoter system of gene expression in C. elegans. *Dev Biol* 296:537-44

da Graca LS, Zimmerman KK, Mitchell MC, Kozhan-Gorodetska M, Sekiewicz K, et al. 2004. DAF-5 is a Ski oncoprotein homolog that functions in a neuronal TGF beta pathway to regulate C. elegans dauer development. *Development* 131:435-46

Dai H, Hogan C, Gopalakrishnan B, Torres-Vazquez J, Nguyen M, et al. 2000. The zinc finger protein schnurri acts as a Smad partner in mediating the transcriptional response to decapentaplegic. *Dev Biol* 227:373-87

Daly AC, Randall RA, Hill CS. 2008. Transforming growth factor beta-induced Smad1/5 phosphorylation in epithelial cells is mediated by novel receptor complexes and is essential for anchorage-independent growth. *Mol Cell Biol* 28:6889-902

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

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

Dubois CM, Laprise MH, Blanchette F, Gentry LE, Leduc R. 1995. Processing of transforming growth factor beta 1 precursor by human furin convertase. *J Biol Chem* 270:10618-24

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

Estevez M, Attisano L, Wrana JL, Albert PS, Massague J, Riddle DL. 1993. The daf-4 gene encodes a bone morphogenetic protein receptor controlling C. elegans dauer larva development. *Nature* 365:644-9

Feng XH, Derynck R. 1997. A kinase subdomain of transforming growth factor-beta (TGF-beta) type I receptor determines the TGF-beta intracellular signaling specificity. *EMBO J* 16:3912-23

Feng XH, Derynck R. 2005. Specificity and versatility in tgf-beta signaling through Smads. *Annu Rev Cell Dev Biol* 21:659-93

Feng XH, Lin X, Derynck R. 2000. Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-beta. *EMBO J* 19:5178-93

Feng XH, Zhang Y, Wu RY, Derynck R. 1998. The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for smad3 in TGF-beta-induced transcriptional activation. *Genes Dev* 12:2153-63

Foehr ML, Lindy AS, Fairbank RC, Amin NM, Xu M, et al. 2006. An antagonistic role for the C. elegans Schnurri homolog SMA-9 in modulating TGFbeta signaling during mesodermal patterning. *Development* 133:2887-96

Gao S, Steffen J, Laughon A. 2005. Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. *J Biol Chem* 280:36158-64

Garamszegi N, Garamszegi SP, Samavarchi-Tehrani P, Walford E, Schneiderbauer MM, et al. Extracellular matrix-induced transforming growth factor-beta receptor signaling dynamics. *Oncogene* 29:2368-80

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

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* 14:435-51

Gomis RR, Alarcon C, He W, Wang Q, Seoane J, et al. 2006. A FoxO-Smad synexpression group in human keratinocytes. *Proc Natl Acad Sci U S A* 103:12747-52

Goumans MJ, Valdimarsdottir G, Itoh S, Lebrin F, Larsson J, et al. 2003. Activin receptor-like kinase (ALK)1 is an antagonistic mediator of lateral TGFbeta/ALK5 signaling. *Mol Cell* 12:817-28

Gray AM, Mason AJ. 1990. Requirement for activin A and transforming growth factor--beta 1 pro-regions in homodimer assembly. *Science* 247:1328-30

Groppe J, Hinck CS, Samavarchi-Tehrani P, Zubieta C, Schuermann JP, et al. 2008. Cooperative assembly of TGF-beta superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. *Mol Cell* 29:157-68

Gumienny TL, MacNeil LT, Wang H, de Bono M, Wrana JL, Padgett RW. 2007. Glypican LON-2 is a conserved negative regulator of BMP-like signaling in Caenorhabditis elegans. *Curr Biol* 17:159-64

Gunther CV, Riddle DL. 2004. Alternative polyadenylation results in a truncated daf-4 BMP receptor that antagonizes DAF-7-mediated development in Caenorhabditis elegans. *J Biol Chem* 279:39555-64

Hart AH, Willson TA, Wong M, Parker K, Robb L. 2005. Transcriptional regulation of the homeobox gene Mixl1 by TGF-beta and FoxH1. *Biochem Biophys Res Commun* 333:1361-9

Hata A, Seoane J, Lagna G, Montalvo E, Hemmati-Brivanlou A, Massague J. 2000. OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-Smad and Olf signaling pathways. *Cell* 100:229-40

Hayashi H, Abdollah S, Qiu Y, Cai J, Xu YY, et al. 1997. The MAD-related protein Smad7 associates with the TGFbeta receptor and functions as an antagonist of TGFbeta signaling. *Cell* 89:1165-73

Hicar MD, Liu Y, Allen CE, Wu LC. 2001. Structure of the human zinc finger protein HIVEP3: molecular cloning, expression, exon-intron structure, and comparison with paralogous genes HIVEP1 and HIVEP2. *Genomics* 71:89-100

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

Huang T, Kuersten S, Deshpande AM, Spieth J, MacMorris M, Blumenthal T. 2001. Intercistronic region required for polycistronic pre-mRNA processing in Caenorhabditis elegans. *Mol Cell Biol* 21:1111-20

Huangfu D, Anderson KV. 2006. Signaling from Smo to Ci/Gli: conservation and divergence of Hedgehog pathways from Drosophila to vertebrates. *Development* 133:3-14

Hurst HC. 1995. Transcription factors 1: bZIP proteins. *Protein Profile* 2:101-68

Huse M, Muir TW, Xu L, Chen YG, Kuriyan J, Massague J. 2001. The TGF beta receptor activation process: an inhibitor- to substrate-binding switch. *Mol Cell* 8:671-82

Imamura T, Takase M, Nishihara A, Oeda E, Hanai J, et al. 1997. Smad6 inhibits signalling by the TGF-beta superfamily. *Nature* 389:622-6

Inman GJ, Hill CS. 2002. Stoichiometry of active smad-transcription factor complexes on DNA. *J Biol Chem* 277:51008-16

Inman GJ, Nicolas FJ, Hill CS. 2002. Nucleocytoplasmic shuttling of Smads 2, 3, and 4 permits sensing of TGF-beta receptor activity. *Mol Cell* 10:283-94

Inoue T, Thomas JH. 2000. Targets of TGF-beta signaling in Caenorhabditis elegans dauer formation. *Dev Biol* 217:192-204

Jin W, Takagi T, Kanesashi SN, Kurahashi T, Nomura T, et al. 2006. Schnurri-2 controls BMP-dependent adipogenesis via interaction with Smad proteins. *Dev Cell* 10:461-71

Johnstone IL. 2000. Cuticle collagen genes. Expression in Caenorhabditis elegans. *Trends Genet* 16:21-7

Jonk LJ, Itoh S, Heldin CH, ten Dijke P, Kruijer W. 1998. Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-beta, activin, and bone morphogenetic protein-inducible enhancer. *J Biol Chem* 273:21145-52

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

Kamath RS, Martinez-Campos M, Zipperlen P, Fraser AG, Ahringer J. 2001. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. *Genome Biol* 2:RESEARCH0002

Kang Y, Chen CR, Massague J. 2003. A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. *Mol Cell* 11:915-26

Kato Y, Habas R, Katsuyama Y, Naar AM, He X. 2002. A component of the ARC/Mediator complex required for TGF beta/Nodal signalling. *Nature* 418:641-6

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

Korchynskyi O, ten Dijke P. 2002. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. *J Biol Chem* 277:4883-91

Kramer JM. 1994. Genetic analysis of extracellular matrix in C. elegans. *Annu Rev Genet* 28:95-116

Kramer JM, Johnson JJ. 1993. Analysis of mutations in the sqt-1 and rol-6 collagen genes of Caenorhabditis elegans. *Genetics* 135:1035-45

Kramer JM, Johnson JJ, Edgar RS, Basch C, Roberts S. 1988. The sqt-1 gene of C. elegans encodes a collagen critical for organismal morphogenesis. *Cell* 55:555-65

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* 11:984-95

Krishna S, Maduzia LL, Padgett RW. 1999. Specificity of TGFbeta signaling is conferred by distinct type I receptors and their associated SMAD proteins in Caenorhabditis elegans. *Development* 126:251-60

Kunwar PS, Zimmerman S, Bennett JT, Chen Y, Whitman M, Schier AF. 2003. Mixer/Bon and FoxH1/Sur have overlapping and divergent roles in Nodal signaling and mesendoderm induction. *Development* 130:5589-99

Labbe E, Silvestri C, Hoodless PA, Wrana JL, Attisano L. 1998. Smad2 and Smad3 positively and negatively regulate TGF beta-dependent transcription through the forkhead DNA-binding protein FAST2. *Mol Cell* 2:109-20

Lagna G, Hata A, Hemmati-Brivanlou A, Massague J. 1996. Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. *Nature* 383:832-6

Lamesch P, Milstein S, Hao T, Rosenberg J, Li N, et al. 2004. C. elegans ORFeome version 3.1: increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res* 14:2064-9

Levy AD, Yang J, Kramer JM. 1993. Molecular and genetic analyses of the Caenorhabditis elegans dpy-2 and dpy-10 collagen genes: a variety of molecular alterations affect organismal morphology. *Mol Biol Cell* 4:803-17

Liang J, Lints R, Foehr ML, Tokarz R, Yu L, et al. 2003. The Caenorhabditis elegans schnurri homolog sma-9 mediates stage- and cell type-specific responses to DBL-1 BMP-related signaling. *Development* 130:6453-64

Liang J, Yu L, Yin J, Savage-Dunn C. 2007. Transcriptional repressor and activator activities of SMA-9 contribute differentially to BMP-related signaling outputs. *Dev Biol* 305:714-25

Liberati NT, Datto MB, Frederick JP, Shen X, Wong C, et al. 1999. Smads bind directly to the Jun family of AP-1 transcription factors. *Proc Natl Acad Sci U S A* 96:4844-9

Lints R, Emmons SW. 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

Liu F, Ventura F, Doody J, Massague J. 1995. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. *Mol Cell Biol* 15:3479-86

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

Liu X, Wang L, Zhao K, Thompson PR, Hwang Y, et al. 2008. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* 451:846-50

Lundquist EA, Herman RK, Rogalski TM, Mullen GP, Moerman DG, Shaw JE. 1996. The mec-8 gene of C. elegans encodes a protein with two RNA recognition motifs and regulates alternative splicing of unc-52 transcripts. *Development* 122:1601-10

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

Luo S, Kleemann GA, Ashraf JM, Shaw WM, Murphy CT. TGF-beta and insulin signaling regulate reproductive aging via oocyte and germline quality maintenance. *Cell* 143:299-312

Luo S, Shaw WM, Ashraf J, Murphy CT. 2009. TGF-beta Sma/Mab signaling mutations uncouple reproductive aging from somatic aging. *PLoS Genet* 5:e1000789

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* 87:1215-24

Mallo GV, Kurz CL, Couillault C, Pujol N, Granjeaud S, et al. 2002. Inducible antibacterial defense system in C. elegans. *Curr Biol* 12:1209-14

Mango SE. 2001. Stop making nonSense: the C. elegans smg genes. *Trends Genet* 17:646-53

Marty T, Muller B, Basler K, 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* 67:753-91

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

McMahon L, Muriel JM, Roberts B, Quinn M, Johnstone IL. 2003. Two sets of interacting collagens form functionally distinct substructures within a Caenorhabditis elegans extracellular matrix. *Mol Biol Cell* 14:1366-78

Mello CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* 10:3959-70

Miura S, Takeshita T, Asao H, Kimura Y, Murata K, et al. 2000. Hgs (Hrs), a FYVE domain protein, is involved in Smad signaling through cooperation with SARA. *Mol Cell Biol* 20:9346-55

Morita K, Chow KL, 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

Moustakas A, Heldin CH. 2009. The regulation of TGFbeta signal transduction. *Development* 136:3699-714

Muchardt C, Seeler JS, Nirula A, Shurland DL, Gaynor RB. 1992. Regulation of human immunodeficiency virus enhancer function by PRDII-BF1 and c-rel gene products. *J Virol* 66:244-50

Muller B, Hartmann B, Pyrowolakis G, Affolter M, Basler K. 2003. Conversion of an extracellular Dpp/BMP morphogen gradient into an inverse transcriptional gradient. *Cell* 113:221-33

Naar AM, Lemon BD, Tjian R. 2001. Transcriptional coactivator complexes. *Annu Rev Biochem* 70:475-501

Nagamatsu Y, Ohshima Y. 2004. Mechanisms for the control of body size by a G-kinase and a downstream TGFbeta signal pathway in Caenorhabditis elegans. *Genes Cells* 9:39-47

Nakao A, Afrakhte M, Moren A, Nakayama T, Christian JL, et al. 1997. Identification of Smad7, a TGFbeta-inducible antagonist of TGF-beta signalling. *Nature* 389:631-5

Nguyen CQ, Hall DH, Yang Y, Fitch DH. 1999. Morphogenesis of the Caenorhabditis elegans male tail tip. *Dev Biol* 207:86-106

Nystrom J, Shen ZZ, Aili M, Flemming AJ, Leroi A, 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

O'Connor MB, Umulis D, Othmer HG, Blair SS. 2006. Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. *Development* 133:183-93

Oukka M, Kim ST, Lugo G, Sun J, Wu LC, Glimcher LH. 2002. A mammalian homolog of Drosophila schnurri, KRC, regulates TNF receptor-driven responses and interacts with TRAF2. *Mol Cell* 9:121-31

Pardali K, Kurisaki A, Moren A, ten Dijke P, Kardassis D, Moustakas A. 2000. Role of Smad proteins and transcription factor Sp1 in p21(Waf1/Cip1) regulation by transforming growth factor-beta. *J Biol Chem* 275:29244-56

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* 11:2679-90

Patterson GI, Padgett RW. 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* 7:229-40

Rachez C, Freedman LP. 2001. Mediator complexes and transcription. *Curr Opin Cell Biol* 13:274-80

Randall RA, Germain S, Inman GJ, Bates PA, Hill CS. 2002. Different Smad2 partners bind a common hydrophobic pocket in Smad2 via a defined proline-rich motif. *EMBO J* 21:145-56

Randall RA, Howell M, Page CS, Daly A, Bates PA, Hill CS. 2004. Recognition of phosphorylated-Smad2-containing complexes by a novel Smad interaction motif. *Mol Cell Biol* 24:1106-21

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* 274:1389-91

Roberts AB, Sporn MB. 1993. Physiological actions and clinical applications of transforming growth factor-beta (TGF-beta). *Growth Factors* 8:1-9

Roberts AF, Gumienny TL, Gleason RJ, Wang H, Padgett RW. Regulation of genes affecting body size and innate immunity by the DBL-1/BMP-like pathway in Caenorhabditis elegans. *BMC Dev Biol* 10:61

Sano Y, Harada J, Tashiro S, Gotoh-Mandeville R, Maekawa T, Ishii S. 1999. ATF-2 is a common nuclear target of Smad and TAK1 pathways in transforming growth factor-beta signaling. *J Biol Chem* 274:8949-57

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

Savage-Dunn C. 2005. TGF-beta signaling. *WormBook*:1-12

Savage-Dunn C, Maduzia LL, Zimmerman CM, Roberts AF, Cohen S, et al. 2003. Genetic screen for small body size mutants in C. elegans reveals many TGFbeta pathway components. *Genesis* 35:239-47

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

Savage C, Das P, Finelli AL, Townsend SR, Sun CY, et al. 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

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* 139:1347-58

Selvamurugan N, Kwok S, Partridge NC. 2004. Smad3 interacts with JunB and Cbfa1/Runx2 for transforming growth factor-beta1-stimulated collagenase-3 expression in human breast cancer cells. *J Biol Chem* 279:27764-73

Seoane J, Le HV, Shen L, Anderson SA, Massague J. 2004. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* 117:211-23

Sherwood DR, Butler JA, Kramer JM, Sternberg PW. 2005. FOS-1 promotes basement-membrane removal during anchor-cell invasion in C. elegans. *Cell* 121:951-62

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

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

Shimmi O, Umulis D, Othmer H, O'Connor MB. 2005. Facilitated transport of a Dpp/Scw heterodimer by Sog/Tsg leads to robust patterning of the Drosophila blastoderm embryo. *Cell* 120:873-86

Shukla A, Malik M, Cataisson C, Ho Y, Friesen T, et al. 2009. TGF-beta signalling is regulated by Schnurri-2-dependent nuclear translocation of CLIC4 and consequent stabilization of phospho-Smad2 and 3. *Nat Cell Biol* 11:777-84

Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, et al. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* 107:465-76

Souchelnytskyi S, ten Dijke P, Miyazono K, Heldin CH. 1996. Phosphorylation of Ser165 in TGF-beta type I receptor modulates TGF-beta1-induced cellular responses. *EMBO J* 15:6231-40

Spike CA, Davies AG, Shaw JE, Herman RK. 2002. MEC-8 regulates alternative splicing of unc-52 transcripts in C. elegans hypodermal cells. *Development* 129:4999-5008

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

Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev Biol* 56:110-56

Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol* 100:64-119

Suzuki Y, Morris GA, Han M, Wood WB. 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

Suzuki Y, Yandell MD, Roy PJ, Krishna S, Savage-Dunn C, et al. 1999. A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in Caenorhabditis elegans. *Development* 126:241-50

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

Thein MC, McCormack G, Winter AD, Johnstone IL, Shoemaker CB, Page AP. 2003. Caenorhabditis elegans exoskeleton collagen COL-19: an adult-specific marker for collagen modification and assembly, and the analysis of organismal morphology. *Dev Dyn* 226:523-39

Tian C, Sen D, Shi H, Foehr ML, Plavskin Y, et al. The RGM protein DRAG-1 positively regulates a BMP-like signaling pathway in Caenorhabditis elegans. *Development* 137:2375-84

Torres-Vazquez J, Park S, Warrior R, Arora K. 2001. The transcription factor Schnurri plays a dual role in mediating Dpp signaling during embryogenesis. *Development* 128:1657-70

van der Keyl H, Kim H, Espey R, Oke CV, Edwards MK. 1994. Caenorhabditis elegans sqt-3 mutants have mutations in the col-1 collagen gene. *Dev Dyn* 201:86-94

von Mende N, Bird DM, Albert PS, Riddle DL. 1988. dpy-13: a nematode collagen gene that affects body shape. *Cell* 55:567-76

Wang G, Long J, Matsuura I, He D, Liu F. 2005. The Smad3 linker region contains a transcriptional activation domain. *Biochem J* 386:29-34

Wang J, Tokarz R, Savage-Dunn C. 2002. The expression of TGFbeta signal transducers in the hypodermis regulates body size in C. elegans. *Development* 129:4989-98

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* 272:11444-51

Wotton D, Knoepfler PS, Laherty CD, Eisenman RN, Massague J. 2001. The Smad transcriptional corepressor TGIF recruits mSin3. *Cell Growth Differ* 12:457-63

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

Wrana JL, Attisano L, Carcamo J, Zentella A, Doody J, et al. 1992. TGF beta signals through a heteromeric protein kinase receptor complex. *Cell* 71:1003-14

Wu JW, Krawitz AR, Chai J, Li W, Zhang F, et al. 2002. Structural mechanism of Smad4 recognition by the nuclear oncoprotein Ski: insights on Ski-mediated repression of TGF-beta signaling. *Cell* 111:357-67

Wu LC. 2002. ZAS: C2H2 zinc finger proteins involved in growth and development. *Gene Expr* 10:137-52

Wu RY, Zhang Y, Feng XH, Derynck R. 1997. Heteromeric and homomeric interactions correlate with signaling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Mol Cell Biol* 17:2521-8

Xu M, Kirov N, Rushlow C. 2005. Peak levels of BMP in the Drosophila embryo control target genes by a feed-forward mechanism. *Development* 132:1637-47

Yao LC, Blitz IL, Peiffer DA, Phin S, Wang Y, et al. 2006. Schnurri transcription factors from Drosophila and vertebrates can mediate Bmp signaling through a phylogenetically conserved mechanism. *Development* 133:4025-34

Yeo C, Whitman M. 2001. Nodal signals to Smads through Cripto-dependent and Cripto-independent mechanisms. *Mol Cell* 7:949-57

Zawel L, Dai JL, Buckhaults P, Zhou S, Kinzler KW, et al. 1998. Human Smad3 and Smad4 are sequence-specific transcription activators. *Mol Cell* 1:611-7

Zhang Y, Feng X, We R, Derynck R. 1996. Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. *Nature* 383:168-72

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

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* 2:121-7

Zugasti O, Rajan J, Kuwabara PE. 2005. The function and expansion of the Patched- and Hedgehog-related homologs in C. elegans. Genome Res 15:1402-10