The *C. elegans* SMOC-1 protein acts cell non-autonomously to promote bone morphogenetic protein signaling

Melisa S. DeGroot, Herong Shi, Alice Eastman, Alexandra N. McKillop, Jun Liu

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

5 **Running Title**: SMOC-1 promotes BMP signaling

6 **Key words**: BMP, SMOC-1, SPARC, LON-2, glypican, body size

7 **Corresponding author**: Jun Liu, Department of Molecular Biology and Genetics, 439

8 Biotechnology Building, Cornell University, Ithaca, NY 14853.

9 Tel: 607-257-2863. Fax: 607-255-6249. Email: JL53@cornell.edu

ABSTRACT (219 words)

Bone morphogenetic protein (BMP) signaling regulates many different developmental and homeostatic processes in metazoans. The BMP pathway is conserved in *Caenorhabditis elegans*, and is known to regulate body size and mesoderm development. We have identified the *C. elegans smoc-1* (Secreted MOdular Calcium binding protein-1) gene as a new player in the BMP pathway. *smoc-1(0)* null mutants have a small body size, while overexpression of *smoc-1* led to a long body size and increased expression of the RAD-SMAD BMP reporter, suggesting that SMOC-1 acts as a positive modulator of BMP signaling. Using double mutant analysis, we showed that SMOC-1 antagonizes the function of the glypican LON-2 and acts through the BMP ligand DBL-1 to regulate BMP signaling. Moreover, SMOC-1 appears to specifically regulate BMP signaling without significant involvement in a TGFβ-like pathway that regulates dauer development. We found that *smoc-1* is expressed in multiple tissues, including cells of the pharynx, intestine, and posterior hypodermis, and that the expression of *smoc-1* in the intestine is positively regulated by BMP signaling. We further established that SMOC-1 functions cell non-autonomously to regulate body size. Human SMOC1 and SMOC2 can each partially rescue the *smoc-1(0)* mutant phenotype, suggesting that SMOC-1's function in modulating BMP signaling is evolutionarily conserved. Together, our findings highlight a conserved role of SMOC proteins in modulating BMP signaling in metazoans.

ARTICLE SUMMARY (100 words)

BMP signaling is critical for development and homeostasis in metazoans, and is under tight regulation. We report the identification and characterization of a Secreted MOdular Calcium binding protein SMOC-1 as a positive modulator of BMP signaling in *C. elegans*. We established that SMOC-1 antagonizes the function of LON-2/glypican and acts through the DBL-1/BMP ligand to promote BMP signaling. We identified *smoc-1*-expressing cells, and demonstrated that SMOC-1 acts cell non-autonomously and in a positive feedback loop to regulate BMP signaling. We also provide evidence suggesting that the function of SMOC proteins in the BMP pathway is conserved from worms to humans.

INTRODUCTION

Bone morphogenetic proteins (BMPs) are highly conserved signaling molecules that mediate cell-cell communication. The BMP signaling cascade is initiated when the BMP ligands bind to the membrane-bound receptor kinases, upon which the type-II receptor phosphorylates the type-I receptors. The signaling cascade is then transduced within the receiving cell as the receptor-associated Smads (R-Smads) are activated via phosphorylation by the type-I receptor. Activated R-Smads complex together with common mediator Smads (co-Smads) and other transcription factors to regulate transcription of downstream genes (KATAGIRI AND WATABE 2016). BMPs regulate fundamental cellular processes, including cell migration, cell proliferation, cell fate specification, and cell death throughout metazoan development (WANG *et al.* 2014). Tight regulation of BMP signaling in time, space, magnitude, and duration is therefore important for proper developmental outcomes. Mis-regulation of BMP signaling can cause a variety of disorders in humans (BRAZIL *et al.* 2015; SALAZAR *et al.* 2016; WU *et al.* 2016) . Previous studies have demonstrated that BMP signaling can be regulated at many levels, both extracellularly and intracellularly (BRAGDON *et al.* 2011; LOWERY *et al.* 2016; SEDLMEIER AND SLEEMAN 2017). The nematode *C. elegans* provides a useful system for identifying factors that modulate the BMP pathway.

The BMP pathway in *C. elegans* is comprised of evolutionarily conserved core components including the ligand (DBL-1/BMP), the type I and type II receptors (SMA-6/RI and DAF-4/RII), the R-Smads (SMA-2 and SMA-3), and the co-Smad (SMA-4) (ESTEVEZ *et al.* 1993; SAVAGE *et al.* 1996; KRISHNA *et al.* 1999; MORITA *et al.* 1999; SUZUKI *et al.* 1999; MORITA *et al.* 2002) (Figure 1A). Unlike in *Drosophila* and vertebrates, BMP signaling is not

essential for viability in *C. elegans*, yet it regulates multiple processes, including body size, male tail development, and mesoderm patterning (GUMIENNY AND SAVAGE-DUNN 2013; SAVAGE-DUNN AND PADGETT 2017). The BMP ligand DBL-1 is expressed in the ventral nerve cord (SUZUKI *et al.* 1999), and it activates the pathway in the hypodermis to regulate body size (YOSHIDA *et al.* 2001; WANG *et al.* 2002). Reduced BMP signaling causes a small (Sma) body size, while increased BMP signaling leads to a long (Lon) body size (MORITA *et al.* 1999; SUZUKI *et al.* 1999; MORITA *et al.* 2002). BMP signaling also regulates the development of the postembryonic mesoderm lineage, the M lineage. We have shown that mutations in the BMP pathway specifically suppress the M lineage dorsoventral patterning defects caused by mutations in *sma-9*, which encodes the *C. elegans* zinc finger protein Schnurri (LIANG *et al.* 2003; FOEHR *et al.* 2006). Specifically, mutations in *sma-9* result in the loss of the two M-derived coelomocytes (CCs), while BMP pathway mutations can restore these two CCs in the *sma-9(0)* mutant background (FOEHR *et al.* 2006; LIU *et al.* 2015; WANG *et al.* 2017) (Figure 1B,C). Using this suppression of *sma-9(0)* M-lineage defect (Susm) assay, we have identified multiple evolutionarily conserved modulators of BMP signaling. These include the RGM protein DRAG-1 (TIAN *et al.* 2010), the neogenin homolog UNC-40 (TIAN *et al.* 2013), the ADAM10 protein SUP-17 (WANG *et al.* 2017), and three tetraspanins, TSP-21, TSP-12 and TSP-14 (LIU *et al.* 2015; WANG *et al.* 2017).

In this study, we report the identification and characterization of a new BMP modulator, which we have named SMOC-1. SMOC-1 is predicted to be a secreted protein that contains a thyroglobulin-like (TY) domain and an extracellular calcium-binding (EC) motif. We show here that SMOC-1 acts as a positive modulator of BMP signaling in *C. elegans*. We further demonstrate that SMOC-1 acts upstream of the ligand to regulate body size. We identified *smoc-*

1-expressing cells, and demonstrated that SMOC-1 acts cell non-autonomously to regulate BMP signaling. Finally, we provide evidence that the function of SMOC proteins in the BMP pathway is conserved from worms to humans.

MATERIALS & METHODS

C. elegans strains

All strains were maintained at 20°C using standard culture conditions (BRENNER 1974) unless otherwise specified. Table 1 lists all the strains used in this study.

Plasmid constructs and transgenic lines

All plasmid constructs used in this study are listed in Table 2. The *smoc-1* open reading frame was amplified from the Vidal RNAi library (RUAL *et al.* 2004). Subsequent sequencing of the clone revealed the presence of a point mutation (S103P, Figure 2D), changing amino acid 103 from serine (TCC) to proline (CCC). Site directed mutagenesis was used to fix this point mutation. Plasmids containing the human *SMOC1* and *SMOC2* cDNAs were purchased from PlasmID, the DNA resource core at Harvard Medical School.

Transgenic strains were generated using the plasmid pRF4 *(rol-6(su1006))*, pCFJ90 (*myo-2p*::*mCherry*::*unc-54 3'UTR)*, or pJKL724 (*myo-3p*::*mCherry*::*unc-54 3' UTR)* as a co-injection marker. Two transgenic lines with the best transmission efficiency were analyzed for each plasmid of interest. Integrated transgenic lines either overexpressing *smoc-1* (*jjIs5119*) or carrying the *smoc-1* transcriptional reporter (*jjIs4688* and *jjIs4694*) were generated using gamma-irradiation, followed by three rounds of outcrossing with N2 worms.

106

107 **Protein sequence alignment**

108 Sequences where taken from Genbank (*C. elegans* SMOC-1 (T04F3.2), 179609; *C.*
109 *remanei* CRE_26999, 9815068; *C. briggsae* CBG23276, 8578577; *D. melanogaster* Pent/Magu,
110 44850; *H. sapiens* SMOC1, 64093; *H. sapiens* SMOC2, 64094) or Wormbase (*C. brenneri*
111 CBN20462; *C. japonica* CJA07338*; P. pacificus* PPA34808). TY and EC domains in SMOC
112 proteins were predicted by Interpro (FINN *et al.* 2017). Domains were aligned using M-COFFEE
113 Multiple Sequence Alignment (MSA) tool on the T-COFFEE server (version 11.00.d625267,
114 (WALLACE *et al.* 2006)). ALN files were processed to produce alignment images using
115 BOXSHADE.

116

117 **Microscopy**

118 Epifluorescence and differential interference contrast (DIC) microscopy were conducted on
119 a Leica DMRA2 compound microscope equipped with a Hamamatsu Orca-ER camera using the
120 iVision software (Biovision Technology, Inc.). Subsequent image analysis was performed using
121 Fiji (SCHINDELIN *et al.* 2012). RAD-SMAD reporter assay was carried out as previously
122 described (TIAN *et al.* 2013).

123

124 **Body size measurements**

125 Body size measurement assays were conducted as previously described (TIAN *et al.* 2013).
126 Hermaphrodite worms were imaged at the L4.3 stage based on vulva development (MOK *et al.*
127 2015). Body sizes were measured from images using the segmented line tool of Fiji. An

ANOVA and a Tukey HSD were conducted to test for differences in body size between genotypes using R (R CORE TEAM 2015).

Suppression of *sma-9(0)* M-lineage defect (Susm) assay

For the Susm assay, worms were grown at 20°C and then the number of animals with 4 CCs and 6 CCs were tallied across three to seven plates for each genotype. For the Susm rescue experiments, we generated general linear models (GLMs) with binomial errors, and a logit link function designating transgene as the explanatory function to test for differences between transgenic and non-transgenic groups within a line.

Dauer formation assay

Dauer formation assay was conducted under non-dauer-inducing conditions as previously described (VOWELS AND THOMAS 1992). Ten adult hermaphrodites were placed on a six centimeter NGM plate (five plates per strain at each temperature) and allowed to lay eggs for less than eight hours. Adults were removed and plates were placed at the test temperature. When non-dauer worms became young adults, the numbers of dauer and non-dauer worms on each plate were scored. Using R, we tested for differences in dauer formation between genotypes using an ANOVA followed by a TukeyHSD.

Data availability statement

Strains and plasmids are available upon request. The authors affirm that all data necessary for confirming the conclusions of the article are present within the article, figures, and tables.

RESULTS

Mutations in T04F3.2 suppress the mesoderm defects of *sma-9(0)* mutants

In a previous *sma-9* suppressor screen, we uncovered a novel complementation group named *susm-1* that includes three alleles, *jj65*, *jj85* and *jj180* (LIU *et al.* 2015) (Table 3), which suppressed the *sma-9(0)* M lineage defect at high penetrance. Whole genome sequencing (WGS) of the three alleles identified molecular lesions in the uncharacterized gene T04F3.2: *jj65* and *jj85* are missense mutations C210Y and E105K, respectively, while *jj180* is a nonsense mutation Q180Stop (Figure 2A,B). To confirm that T04F3.2 is the corresponding gene for this complementation group, we obtained two deletion alleles that delete most of the coding region of T04F3.2, *tm7000* and *tm7125* (Figure 2A), and found that both alleles suppressed the *sma-9(0)* M lineage defect to near 100% (Table 3, Figure 1C). Pairwise complementation tests between *tm7000* and *jj65, jj85* or *jj180*, showed that *tm7000* failed to complement all three alleles in their suppression of the *sma-9(0)* M lineage defect (Table 3). Subsequent *sma-9(0)* suppressor screens conducted in the lab identified three additional alleles of this complementation group, *jj109, jj115,* and *jj139*. WGS followed by Sanger sequencing showed that all three alleles contain nonsense mutations in T04F3.2: W13Stop for both *jj115* and *jj139*, and W176Stop for *jj109* (Figure 2A,B). Finally, a transgene containing the T04F3.2 genomic region including 2kb upstream sequences, the entire coding region with introns, and 2kb downstream sequences rescued the *sma-9(0)* suppression phenotype of *tm7125* mutants (Table 3). Collectively, these results demonstrated that T04F3.2 is the corresponding gene for the *susm-1* locus. The nature of the molecular lesions in *tm7000*, *tm7125*, *jj109*, *jj115*, *jj139,* and *jj180*, the near 100% penetrance of their Susm phenotypes, and their similar body size phenotypes (see below),

174 suggest that all of these alleles are null alleles. For ease of genotyping, most of our subsequent
175 analysis was carried out using the *tm7125* allele.
176
177 **T04F3.2 encodes a predicted Secreted MOdular Calcium-binding protein SMOC-1**
178 T04F3.2 is predicted to encode a protein of 260 amino acids. It contains a predicted signal
179 peptide (SP), a thyroglobulin type I-like repeat (TY), and a secreted protein acidic and rich in
180 cysteine (SPARC) extracellular calcium (EC) binding region (Figure 2). The EC domain is
181 predicted to contain a pair of helix-loop-helix EF hand calcium-binding motifs (HOHENESTER *et*
182 *al.* 1996; VANNAHME *et al.* 2002). The predicted T04F3.2 protein is most similar to the human
183 secreted modular calcium-binding proteins SMOC1 and SMOC2 (VANNAHME *et al.* 2002;
184 VANNAHME *et al.* 2003), and the *Drosophila melanogaster* SMOC homolog Pentagone/Magu
185 (VUILLEUMIER *et al.* 2010). A BLAST search against the *C. elegans* genome showed that
186 T04F3.2 is the only SMOC homolog. We have, therefore, named this gene *smoc-1* and its
187 corresponding protein SMOC-1.
188 SMOC proteins are matricellular proteins that are in the same family as SPARC/BM-
189 40/osteonectin (BRADSHAW 2012). The domain arrangement of SMOC proteins varies across
190 species. The *C. elegans* SMOC-1 protein is predicted to have one TY domain, one EC domain,
191 and completely lack the follistatin (FS) domain that is present in other SMOC proteins (Figure
192 2C). Within the TY domain, SMOC-1 shares about 30% amino acid identity and 50% similarity
193 with human SMOC1 and SMOC2, and contains a CWCV tetrapeptide sequence and an
194 additional four conserved cysteines that are characteristic of the TY domain (Figure 2D). The EC
195 domain of SMOC-1 shares about 25% amino acid identity and 45% similarity with those of the

196 human SMOC proteins. Among the conserved residues in the EC domain are four cysteines
197 thought to be involved in disulfide bond formation (BUSCH *et al.* 2000).

198 The locations of the molecular lesions in our *smoc-1* mutant alleles suggest that both the
199 TY domain and the EC domain are important for SMOC-1 function. *jj85* is a mutation in the TY
200 domain, changing amino acid 105 from a glutamic acid to a lysine (E105K, Figure 2B,D).
201 Although the change appears to make this residue more similar to its counterpart (arginine or
202 lysine) in the fly and human SMOC proteins (Figure 2D), we noted that E105 is conserved in
203 multiple nematode species (Figure 2F). We also obtained a *smoc-1* cDNA clone that has a single
204 base mutation changing amino acid 103 from a conserved serine to proline (Figure 2D). This
205 mutant *smoc-1* cDNA (S103P) failed to rescue the *smoc-1(0)* Susm phenotype, while the wild-
206 type (WT) *smoc-1* cDNA under the same regulatory elements successfully rescued the *smoc-1(0)*
207 Susm phenotype (Table 3), again highlighting the importance of the TY domain for SMOC-1
208 function. Similarly, the EC domain is also critical for SMOC-1 function, because a change of the
209 conserved cysteine residue at amino acid 210 to tyrosine (C210Y) in *jj65* significantly
210 compromised the function of SMOC-1 (Figure 2B,E, Table 3).
211
212 **SMOC-1 functions within the BMP pathway to positively regulate BMP signaling**
213 We have previously shown that mutations in BMP pathway components specifically
214 suppress the *sma-9(0)* M lineage defect (FOEHR *et al.* 2006; LIU *et al.* 2015). The highly
215 penetrant Susm phenotype of multiple *smoc-1* alleles suggests that SMOC-1 may function in the
216 BMP pathway. BMP pathway mutants are known to exhibit altered body sizes (SAVAGE-DUNN
217 AND PADGETT 2017). We measured the body sizes of *smoc-1* single mutant animals and found
218 that they all have a reproducibly smaller body size (~95%) compared to WT animals at the same

219 developmental stage (Figure 3A,B,D). This smaller body size can be rescued by a WT *smoc-1*
220 transgene (Figure 3D). Moreover, transgenic *smoc-1* mutant animals carrying this transgene are
221 significantly longer than WT animals (Figure 3D). This is likely due to the repetitive nature of
222 the transgene generated using standard *C. elegans* transgenic approaches, which often results in
223 over-expression of the transgene (MELLO *et al.* 1991). We have subsequently integrated the WT
224 *smoc-1* transgene in the WT background (*jjIs5119*, Table 1). Again, *jjIs5119* (which we have
225 referred to as *smoc-1(OE)*) animals are significantly longer than WT animals (Figure 4B). Thus,
226 *smoc-1* appears to function in a dose-dependent manner to positively regulate body size.
227 To determine whether *smoc-1* functions within the BMP pathway to regulate body size, we
228 generated double mutants between *smoc-1(tm7125)* and null mutations in various BMP pathway
229 components, and measured their body lengths. As shown in Figure 4A, *dbl-1(ok3749) smoc-*
230 *1(tm7125)* double mutants were as small as *dbl-1(ok3749)* single mutants. Similarly, *sma-3(jj3);*
231 *smoc-1(tm7125)* and *sma-6(jj1); smoc-1(tm7125)* double mutants were as small as *sma-3(jj3)*
232 and *sma-6(jj1)* single mutants, respectively. These observations indicate that *smoc-1* functions
233 within the BMP pathway, rather than in a parallel pathway, to regulate body size.
234 In addition to body size, BMP pathway mutants also exhibit male tail defects and the
235 mutant males cannot mate (SAVAGE *et al.* 1996; KRISHNA *et al.* 1999; SUZUKI *et al.* 1999). We
236 generated *smoc-1(tm7125)* males and found that they mated well with WT hermaphrodites to
237 produce cross progeny, suggesting that *smoc-1(tm7125)* males do not have severe male tail
238 patterning defects. This is not surprising as previous studies have demonstrated that male tail
239 development is not affected when there is a partial reduction of BMP signaling (KRISHNA *et al.*
240 1999).

We also examined the expression of the RAD-SMAD reporter, which we have previously shown to serve as a direct readout of BMP signaling (TIAN *et al.* 2010). While *smoc-1* null mutants did not exhibit significant changes in the expression of the RAD-SMAD reporter (data now shown), the *smoc-1(OE)* lines showed a significant increase in the level of RAD-SMAD reporter expression (Figure 4C,D). We reasoned that the change of RAD-SMAD reporter expression in *smoc-1(0)* mutants may be too small to detect given that *smoc-1(0)* mutants only exhibit about 5% reduction in body size compared to WT animals (see above). Nevertheless, our findings are consistent with SMOC-1 acting in the BMP pathway to positively promote BMP signaling.

SMOC-1 functions through the BMP ligand to promote BMP signaling in regulating body size

The long body size phenotype caused by *smoc-1* overexpression provided us with a useful tool to determine where in the BMP signaling pathway SMOC-1 functions. We conducted genetic epistasis analysis by generating double mutants between *smoc-1(OE)* and null mutations in core components of the BMP pathway that are known to cause a small body size. As shown in Figure 4B, *smoc-1(OE); dbl-1(ok3749)* double mutants and *smoc-1(OE); sma-3(tm4625)* double mutants are as small as *dbl-1(ok3749)* and *sma-3(tm4625)* single mutants, respectively. These results provide further support to the conclusion that SMOC-1 functions within the BMP pathway to regulate body size. More importantly, our genetic epistasis results demonstrate that SMOC-1 functions upstream of and is dependent on the function of the BMP ligand DBL-1 to regulate body size.

SMOC-1 antagonizes the function of LON-2/glypican to modulate BMP signaling in regulating body size

Previous studies have shown that the glypican LON-2 functions upstream of DBL-1/BMP and acts as a negative regulator of BMP signaling (GUMIENNY *et al.* 2007). We performed double mutant analysis and dissected the relationship between SMOC-1 and LON-2/glypican. We first measured the body length of double null mutants between *smoc*-1 and *lon-2*. As shown in Figure 5A, *smoc-1(tm7125); lon-2(e678)* double null mutants exhibited an intermediate body size compared to either single null mutant. In particular, the body size of *smoc-1(tm7125); lon-2(e678)* double mutants is similar to that of WT animals. These observations suggest that SMOC-1 and LON-2/glypican antagonize each other in regulating body size. Interestingly, *smoc-1(OE); lon-2(e678)* worms are longer than either *smoc-1(OE)* animals or *lon-2(e678)* single mutants (Figure 5B). Thus, over-expressing *smoc-1* is capable of further increasing the body size of worms that completely lack LON-2/glypican. Taken together, our genetic analysis between *lon-2* and *smoc-1* suggests that SMOC-1 antagonizes the function of LON-2/glypican in regulating body size, and that SMOC-1 also has LON-2/glypican-independent function(s) in promoting BMP signaling.

SMOC-1 does not play a major role in the TGFβ-like dauer pathway

In addition to the BMP pathway, *C. elegans* has a TGFβ-like signaling pathway that regulates dauer development (SAVAGE-DUNN AND PADGETT 2017). To determine if SMOC-1 plays a role in the TGFβ-like dauer pathway, we first assayed dauer formation of worms with different levels of *smoc-1* expression. *smoc-1(tm7125)* and *smoc-1(OE)* single mutant worms did not exhibit any constitutive or defective dauer formation phenotype at any of the temperatures

tested (Table 4, data not shown), suggesting that SMOC-1 does not play a major role in the TGFβ-like dauer pathway. Next, we generated double mutant worms carrying both *smoc-1(tm7125)* and mutations in the TGFβ ligand DAF-7/TGFβ or the type 1 receptor DAF-1/RI (GEORGI *et al.* 1990; REN *et al.* 1996), and examined them for the constitutive dauer formation (Daf-c) phenotype (Table 1). While *smoc-1(tm7125)* partially suppressed the Daf-c phenotype of *daf-7(e1372)* at 20°C, a similar trend was not observed at either 15°C or at 25°C. Similarly, *smoc-1(tm7125)* did not exhibit any consistent suppression or enhancement of the Daf-c phenotype of two *daf-1* mutant alleles (Table 4). These results suggest that SMOC-1 does not play a major role in the TGFβ-like dauer pathway, although we cannot rule out a minor buffering function of SMOC-1 in this pathway.

Because of the genetic interaction that we observed between *smoc-1* and *lon-2*, we also tested whether LON-2/glypican plays a role in the TGFβ-like dauer pathway by performing similar double mutant analysis as described for *smoc-1*. At 20°C, *lon-2(e678)* showed partial suppression of the Daf-c phenotype of *daf-7(e1372)* (Table 5), but a similar trend was not observed at 15°C or at 25°C (Table 5). As seen with *smoc-1(tm7125)*, *lon-2(e678)* also did not consistently enhance or suppress the Daf-c phenotype of a TGFβ receptor mutation, *daf-1(m213)*. Thus, like SMOC-1, LON-2 does not appear to play a major role, but may play a minor modulatory role, in the TGFβ dauer pathway.

smoc-1 is expressed in the pharynx, intestine, and posterior hypodermis

Since *smoc-1* is predicted to encode a secreted protein, we first attempted to identify the cells that express *smoc-1*. As described above, a *smoc-1* genomic fragment containing 2kb upstream sequences, the entire coding region with introns, and 2kb downstream sequences

(pJKL1128, Table 2) can rescue the Susm and body size phenotypes of *smoc-1(0)* mutants (Figure 6A, Table 3). The same promoter element driving the *smoc-1* cDNA with its own 3'UTR or with the *unc-54* 3'UTR rescued both the small body size and the Susm phenotypes of *smoc-1(0)* mutants (Figure 6A; Table 3), suggesting that the regulatory elements required for SMOC-1 function in BMP signaling reside in the 2kb upstream sequences. We therefore generated a transcriptional reporter *pJKL1139[smoc-1 2kb promoter::4xnls::gfp::unc-54 3'UTR]* (Table 2). We also generated two additional transcriptional reporters using 5kb *smoc-1* upstream sequences (*pJKL1201[smoc-1 5kb promoter::4xnls::gfp::unc-54 3'UTR]* and *pJKL1202[smoc-1 5kb promoter::4xnls::gfp::2kb smoc-1 3'UTR]*, Table 2). All three reporters showed similar expression patterns in transgenic animals. We therefore focused on *pJKL1139[smoc-1 2kb promoter::4xnls::gfp::unc-54 3'UTR]* and generated integrated transgenic lines carrying this reporter (*jjIs4688* and *jjIs4694*, Table 1) for further analysis.

The integrated *smoc-1* transcriptional reporter showed strong GFP expression. GFP was first detectable in several cells located in the anterior of bean stage embryos (Fig 6F). In the developing larvae, GFP is expressed in cells of the pharynx, the intestine and the posterior hypodermis (Fig 6B). Pharyngeal cells expressing *smoc-1p::gfp* include the epithelial cells e2, the marginal cells mc1 and mc2, the M4 neuron, and all six of the pharyngeal/intestinal valve cells (Figure 6C). Cells of the posterior hypodermis expressing *smoc-1p::gfp* include hyp8, hyp9, hyp10, and hyp11 (Fig 6D). Expression in these tissues persisted from the L1 larval stage through adulthood. We noted that while all transgenic animals showed GFP expression in the pharynx and the posterior hypodermis, a small fraction of animals (~8%) did not exhibit GFP expression in all or some of the intestinal cells (Figure 6G). We observed no GFP expression in

any other tissues, including the nerve cord, body wall muscles (BWMs), or the M lineage. Thus, *smoc-1* is expressed in cells of the pharynx, intestine, and posterior hypodermis.

Intestinal expression of *smoc-1* is positively regulated by BMP signaling

We next asked whether *smoc-1* expression is regulated by the BMP pathway or by SMOC-1 itself. We introduced the integrated *smoc-1* transgenic reporter into BMP pathway null mutants, including *sma-3(jj3), sma-6(jj1), lon-2(e678)* and *smoc-1(tm7125)* mutants (Table 1), and examined the expression pattern of the GFP reporter. Intriguingly, while the expression pattern and expression level of the GFP reporter in the pharynx and posterior hypodermis remained relatively constant in all mutant background examined, in *sma-6(jj1)* and *sma-3(jj3)* mutants there was a significant decrease in the percentage of animals that exhibited GFP expression in the intestinal cells and a decrease in the intensity of intestinal GFP expression compared with WT animals (Figure 6F, G). There was also a moderate decrease in the percentage of animals showing intestinal GFP expression in *smoc-1(tm7125)* mutants (Figure 6G). In contrast, nearly 100% of *lon-2(e678)* animals showed bright intestinal GFP expression, as compared to ~92% for WT animals (Figure 6G). Collectively, these results suggest that *smoc-1* expression in the intestinal cells is positively regulated by BMP signaling.

***smoc-1* functions cell non-autonomously to regulate body size and M lineage development**

The *smoc-1* transcriptional reporters identified cells in the pharynx, intestine, and posterior hypodermis as *smoc-1*-expressing cells. To determine in which tissue(s) expression of *smoc-1* is sufficient to regulate BMP signaling, we used a set of promoters to drive *smoc-1* cDNA in a tissue-specific manner, and assayed for rescue of the *smoc-1(tm7125)* mutant phenotypes. Each

355 rescuing construct was introduced into *smoc-1(tm7125)* worms for the body size assay, and into
356 *smoc-1(tm7125); sma-9(cc604)* worms for the Susm assay.
357 As shown in Figure 7A, forced expression of *smoc-1* cDNA specifically within each
358 individual *smoc-1*-expressing tissue (*ifb-2p* for intestinal cells (HUSKEN *et al.* 2008), *myo-2p* for
359 pharyngeal muscles (OKKEMA *et al.* 1993), or *elt-3p* for hypodermal cells (GILLEARD *et al.*
360 1999)) not only rescued the small body size of *smoc-1(tm7125)* mutants, but also made the
361 transgenic worms longer, just like *smoc-1* cDNA under the control of its own promoter. Forced
362 expression of *smoc-1* cDNA in tissues that do not express *smoc-1* (*myo-3p* for BWMs (OKKEMA
363 *et al.* 1993) or *rab-3p* for neurons (NONET *et al.* 1997)) also rescued the small body size of *smoc-*
364 *1(tm7125)* mutants, and made the transgenic worms longer (Figure 7A). An exception is the lack
365 of rescue of the body size phenotype in *smoc-1(tm7125)* mutants upon forced expression of
366 *smoc-1* cDNA in the M lineage using the *hlh-8* promoter (HARFE *et al.* 1998). This could be due
367 to the transient nature of *hlh-8* promoter activity in undifferentiated M lineage cells during larval
368 development (HARFE *et al.* 1998).
369 Similar to the body size rescue results, forced expression of *smoc-1* cDNA in both *smoc-1-*
370 expressing cells (intestine, pharynx, or hypodermis) and cells that do not normally express *smoc-*
371 *1* (BWMs, neurons, or the M lineage) rescued the Susm phenotype of *smoc-1(tm7125)* mutants
372 (Figure 7B), although for reasons currently unknown, the rescuing efficiency appeared lower
373 when *smoc-1* expression was forced in BWMs or neurons (Figure 7B). Taken together, our
374 results demonstrate that SMOC-1 can function cell non-autonomously to regulate both body size
375 and M lineage patterning. This is consistent with SMOC-1 being a putative secreted protein.
376
377 **Human SMOC proteins can partially rescue the *smoc-1(0)* mutant phenotype in *C. elegans***

As described above, SMOC-1 has two human homologs, SMOC1 (hSMOC1) and SMOC2 (hSMOC2). We next asked whether either of the human SMOCs can substitute for SMOC-1 function in *C. elegans*. We first generated plasmids by directly putting the coding region of hSMOC1 or hSMOC2 in between the 2kb *smoc-1* promoter and the *unc-54* 3'UTR (Table 2, Figure 8A), and tested their functionality using the Susm assay. Neither hSMOC1 nor hSMOC2 rescued the Susm phenotype of *smoc-1(tm7125)* worms (Figure 8B). We reasoned that the lack of rescue may be due to differences in the signal peptide between humans and *C. elegans*, causing the proteins to not be properly secreted from cells (TIAN *et al.* 2010). We next generated plasmids expressing chimeric SMOC proteins that have the worm SMOC-1 signal peptide (CelSP) followed by the extracellular region of hSMOC1 or hSMOC2 (Table 2, Figure 8A). Both CelSP::hSMOC1 and CelSP::hSMOC2 partially rescued the Susm phenotype of *smoc-1(tm7125)* mutants (Figure 8B), but failed to rescue the body size phenotype (Figure 8C). Nevertheless, these results demonstrate that CelSP::hSMOC1 and CelSP::hSMOC2 can function to regulate BMP signaling in a *C. elegans* trans-environment and suggest that the function of SMOC proteins in regulating BMP signaling is evolutionarily conserved from worms to humans.

DISCUSSION

In this study, we identified SMOC-1, the sole *C. elegans* SMOC protein that belongs to the SPARC/BM40 family of matricellular proteins, as a key player in the BMP signaling pathway. *smoc-1(0)* mutants have a small body size and suppress the *sma-9(0)* M lineage defect, but *smoc-1(0)* mutants are not as small as null mutants in core components of the BMP pathway (Table 3, Figures 1, 3). These phenotypes resemble those caused by mutations in other modulators of the BMP pathway, such as DRAG-1/RGM (TIAN *et al.* 2010), TSP-21 (LIU *et al.* 2015), or SUP-

17/ADAM10 (WANG *et al.* 2017), and are consistent with a modulatory role for SMOC-1 in the BMP pathway. Over-expression of *smoc-1* led to a significant increase in body size and an increase in RAD-SMAD reporter expression. Moreover, the long body size phenotype caused by *smoc-1(OE)* is completely suppressed by null mutations in the BMP ligand DBL-1 and the R-Smad SMA-3 (Figure 4). Collectively, these findings demonstrate that SMOC-1 functions through the BMP ligand DBL-1 and acts as a positive modulator to promote BMP signaling.

How might SMOC-1 function to promote BMP signaling? Our tissue specific rescue data coupled with the expression pattern of *smoc-1* (Figures 6, 7) showed that SMOC-1 functions cell non-autonomously to regulate BMP signaling. This is consistent with SMOC-1 being a predicted secreted protein. Strikingly, forced expression of *smoc-1* exclusively in pharyngeal muscles is sufficient to rescue both the body size and the Susm phenotype of *smoc-1(0)* mutants (Figure 7). Notably, the M lineage cells, where the Smad proteins function to regulate M lineage development (FOEHR *et al.* 2006), are located in the posterior of a developing larva, distant from the pharynx. Thus SMOC-1 can function over long distances, from a source located far from BMP-receiving cells, to regulate the output of BMP signaling.

The *Drosophila* homolog of SMOC-1, Pent, can also function over long distances to regulate Dpp/BMP signaling in the developing wing imaginal discs (VUILLEUMIER *et al.* 2010). In particular, Pent has been shown to bind to and induce the internalization of the BMP co-receptor Dally/glypican (a heparan sulfate proteoglycans (HSPG)), such that the trapping of Dpp/BMP by Dally is reduced, which in turn promotes the spreading of Dpp/BMP (NORMAN *et al.* 2016). Using a *Xenopus* animal cap transfer assay, Thomas and colleagues (THOMAS *et al.* 2017) showed that *Xenopus* SMOC-1 can also expand the range of BMP signaling by competing with BMP to bind to HSPGs. In *C. elegans*, the glypican homolog LON-2 is a known negative

regulator of BMP signaling, and LON-2 can bind to BMP *in vitro* (GUMIENNY *et al.* 2007). LON-2/glypican has therefore been proposed to negatively regulate BMP signaling by sequestering the DBL-1/BMP ligand. Our genetic analysis between *lon-2(0)* and *smoc-1(0)* null mutations suggests that SMOC-1 antagonizes the function of LON-2 in regulating BMP signaling (Figure 5A). The phenotype of *smoc-1(0); lon-2(0)* double mutants is consistent with a model where SMOC-1 promotes BMP signaling by competing with DBL-1/BMP to bind LON-2/glypican. However, SMOC-1 must have LON-2/glypican-independent function(s), because *smoc-1(OE)* can further increase body size in the absence of LON-2/glypican, as in *smoc-1(OE); lon-2(0)* double mutants shown in Figure 5B.

The molecular mechanism underlying the LON-2/glypican-independent function of SMOC-1 is currently unknown. In addition to LON-2, there are five other HSPG-encoding genes in the *C. elegans* genome: *unc-52* (ROGALSKI *et al.* 1993; HALFTER *et al.* 1998; ACKLEY *et al.* 2001; RHINER *et al.* 2005; HRUS *et al.* 2007). It is possible that in addition to LON-2/glypican, one or multiple of these other HSPGs also functions with SMOC-1 to regulate BMP signaling. Alternatively, SMOC-1 may promote BMP signaling by interacting with other cell surface or extracellular BMP regulators or even with DBL-1/BMP itself to promote BMP signaling. Any LON-2/glypican-independent function of SMOC-1 still requires DBL-1/BMP, because *smoc-1(OE); dbl-1(0)* double mutants are as small as *dbl-1(0)* null mutants. Our model proposing SMOC-1 has dual modes of action to regulate BMP signaling is consistent with structure-function analysis of *Xenopus* SMOC-1 (XSMOC-1), whose EC domains can bind to HSPG and promote BMP spreading, while the TY domains are necessary for XSMOC-1 to inhibit BMP signaling (THOMAS *et al.* 2017). We have shown that both the TY domain and the EC domain in *C. elegans* SMOC-1 are important for its function in BMP signaling, because mutations in either

447 domain disrupt the function of SMOC-1 (Figure 2). Further dissection of the roles of each of
448 these domains at the molecular level will help clarify the mechanisms underlying SMOC-1
449 function in the BMP pathway.

450 In this study, we have shown that in addition to being a positive regulator of BMP
451 signaling, *smoc-1* is also positively regulated by BMP signaling at the transcriptional level
452 (Figure 6). Whether *smoc-1* is directly or indirectly regulated by BMP signaling remains to be
453 determined. Nevertheless, our results suggest a model in which SMOC-1 functions in a positive
454 feedback loop to regulate BMP signaling (Figure 9). Whether or not *smoc-1* expression is
455 directly regulated by BMP signaling is currently unknown.

456 In addition to their roles in regulating BMP signaling, SMOC proteins can also function in
457 other signaling pathways. Pent has been shown to play a role in regulating Wg signaling in the
458 *Drosophila* wing (NORMAN *et al.* 2016). Human SMOC1 can bind to the TGFβ co-receptor
459 endoglin to regulate TGFβ signaling in endothelial cells (AWWAD *et al.* 2015), while SMOC2
460 can potentiate endothelial growth factor or fibroblast growth factor activity to promote
461 angiogenesis in cultured human umbilical vein endothelial cells (HUVECs) (ROCNIK *et al.*
462 2006). Our genetic analysis suggests that SMOC-1 does not play a key role in regulating the
463 TGFβ-like dauer pathway (Table 4). Whether SMOC-1 is involved in other signaling pathways
464 in *C. elegans* is currently unknown.

465 There are two SMOC homologs in mammals. SMOC1 is essential for eye and limb
466 development in mice, and mutations in SMOC1 in humans cause microphthalmia with limb
467 anomalies (MLA) and ophthalmo-acromelic syndrome (OAS) (also known as Waardenburg
468 anophthalmia syndrome (WAS)), both of which affect eye and limb development (OKADA *et al.*
469 2011; RAINGER *et al.* 2011). Mutations in hSMOC2 have also been found to be associated with

defects in dental development (BLOCH-ZUPAN *et al.* 2011; ALFAWAZ *et al.* 2013) and vitiligo (ALKHATEEB *et al.* 2010; BIRLEA *et al.* 2010). QTL mapping in different dog breeds have found that a retrotransposon insertion that disrupts SMOC2 splicing and reduces its expression is associated with canine brachycephaly (MARCHANT *et al.* 2017). In addition, several different types of brain tumors exhibit altered expression of SMOC1 (BRELLIER *et al.* 2011), while SMOC2 is an intestinal stem cell signature gene (MUNOZ *et al.* 2012) that is required for L1-mediated colon cancer progression (SHVAB *et al.* 2016). Notably, BMP signaling is known to play important roles in eye, tooth and limb development, and abnormal BMP signaling can cause cancer (THAWANI *et al.* 2010). Here, we have demonstrated that both hSMOC1 and hSMOC2 can partially rescue the Susm phenotype of *smoc-1(0)* mutants (Figure 8), suggesting that the function of SMOC proteins in regulating BMP signaling is evolutionarily conserved. Future studies on how SMOC-1 functions to regulate BMP signaling in an *in vivo* system such as *C. elegans* may have implications for human health.

ACKNOWLEDGMENTS

We thank Oliver Hobert and Shohei Mitani for plasmids or strains, Florencia Schlamp for advice on statistical analysis using R, Gabriela Rojas for help generating the *smoc-1(OE); RAD-SMAD* strains, and the rest of the Liu lab for critical comments and suggestions. Some strains were obtained from the *C. elegans* Genetics Center, which is funded by NIH Office of 27 Research Infrastructure Programs (P40 OD010440). This study was supported by the National Institutes of Health grants R01 GM066953 and R01 GM103869 to JL. MSD was partially supported by a NIH Predoctoral Training Grant (T32GM007617), a Dean's Excellence Fellowship by the Cornell Graduate School, and an NSF Graduate Research Fellowship (DGE-

493 1650441). AE was a student in the MBG-REU program, which was supported by the NSF
494 (DBI1659534), Department of Molecular Biology and Genetics, Weill Institute of Cell and
495 Molecular Biology, and Division of Nutritional Sciences at Cornell University. ANM was a
496 Hunter R. Rawlings III Presidential Research Scholar at Cornell University.

LITERATURE CITED

Ackley, B. D., J. R. Crew, H. Elamaa, T. Pihlajaniemi, C. J. Kuo *et al.*, 2001 The NC1/endostatin domain of Caenorhabditis elegans type XVIII collagen affects cell migration and axon guidance. J Cell Biol 152**:** 1219-1232.

Alfawaz, S., F. Fong, V. Plagnol, F. S. Wong, J. Fearne *et al.*, 2013 Recessive oligodontia linked to a homozygous loss-of-function mutation in the SMOC2 gene. Arch Oral Biol 58**:** 462-466.

Alkhateeb, A., N. Al-Dain Marzouka and F. Qarqaz, 2010 SMOC2 gene variant and the risk of vitiligo in Jordanian Arabs. Eur J Dermatol 20**:** 701-704.

Awwad, K., J. Hu, L. Shi, N. Mangels, R. Abdel Malik *et al.*, 2015 Role of secreted modular calcium-binding protein 1 (SMOC1) in transforming growth factor beta signalling and angiogenesis. Cardiovasc Res 106**:** 284-294.

Birlea, S. A., K. Gowan, P. R. Fain and R. A. Spritz, 2010 Genome-wide association study of generalized vitiligo in an isolated European founder population identifies SMOC2, in close proximity to IDDM8. J Invest Dermatol 130**:** 798-803.

Bloch-Zupan, A., X. Jamet, C. Etard, V. Laugel, J. Muller *et al.*, 2011 Homozygosity mapping and candidate prioritization identify mutations, missed by whole-exome sequencing, in SMOC2, causing major dental developmental defects. Am J Hum Genet 89**:** 773-781.

Bradshaw, A. D., 2012 Diverse biological functions of the SPARC family of proteins. Int J Biochem Cell Biol 44**:** 480-488.

Bragdon, B., O. Moseychuk, S. Saldanha, D. King, J. Julian *et al.*, 2011 Bone morphogenetic proteins: a critical review. Cell Signal 23**:** 609-620.

Brazil, D. P., R. H. Church, S. Surae, C. Godson and F. Martin, 2015 BMP signalling: agony and antagony in the family. Trends Cell Biol 25**:** 249-264.

Brellier, F., S. Ruggiero, D. Zwolanek, E. Martina, D. Hess *et al.*, 2011 SMOC1 is a tenascin-C interacting protein over-expressed in brain tumors. Matrix Biol 30**:** 225-233.

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

Busch, E., E. Hohenester, R. Timpl, M. Paulsson and P. Maurer, 2000 Calcium affinity, cooperativity, and domain interactions of extracellular EF-hands present in BM-40. J Biol Chem 275**:** 25508-25515.

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

Finn, R. D., T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork *et al.*, 2017 InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 45**:** D190-d199.

Foehr, M. L., A. S. Lindy, R. C. Fairbank, N. M. Amin, M. Xu *et al.*, 2006 An antagonistic role for the C. elegans Schnurri homolog SMA-9 in modulating TGFbeta signaling during mesodermal patterning. Development 133**:** 2887-2896.

Georgi, L. L., P. S. Albert and D. L. Riddle, 1990 daf-1, a C. elegans gene controlling dauer larva development, encodes a novel receptor protein kinase. Cell 61**:** 635-645.

Gilleard, J. S., Y. Shafi, J. D. Barry and J. D. McGhee, 1999 ELT-3: A Caenorhabditis elegans GATA factor expressed in the embryonic epidermis during morphogenesis. Dev Biol 208**:** 265-280.

Gumienny, T. L., L. T. MacNeil, H. Wang, M. de Bono, J. L. Wrana *et al.*, 2007 Glypican LON-2 is a conserved negative regulator of BMP-like signaling in Caenorhabditis elegans. Curr Biol 17**:** 159-164.

Gumienny, T. L., and C. Savage-Dunn, 2013 TGF-beta signaling in C. elegans. WormBook**:** 1-34.

Halfter, W., S. Dong, B. Schurer and G. J. Cole, 1998 Collagen XVIII is a basement membrane heparan sulfate proteoglycan. J Biol Chem 273**:** 25404-25412.

Harfe, B. D., A. Vaz Gomes, C. Kenyon, J. Liu, M. Krause *et al.*, 1998 Analysis of a Caenorhabditis elegans Twist homolog identifies conserved and divergent aspects of mesodermal patterning. Genes Dev 12**:** 2623-2635.

Hohenester, E., P. Maurer, C. Hohenadl, R. Timpl, J. N. Jansonius *et al.*, 1996 Structure of a novel extracellular Ca(2+)-binding module in BM-40. Nat Struct Biol 3**:** 67-73.

Hrus, A., G. Lau, H. Hutter, S. Schenk, J. Ferralli *et al.*, 2007 C. elegans agrin is expressed in pharynx, IL1 neurons and distal tip cells and does not genetically interact with genes involved in synaptogenesis or muscle function. PLoS One 2**:** e731.

Husken, K., T. Wiesenfahrt, C. Abraham, R. Windoffer, O. Bossinger *et al.*, 2008 Maintenance of the intestinal tube in Caenorhabditis elegans: the role of the intermediate filament protein IFC-2. Differentiation 76**:** 881-896.

Katagiri, T., and T. Watabe, 2016 Bone Morphogenetic Proteins. Cold Spring Harb Perspect Biol 8.

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

Liang, J., R. Lints, M. L. Foehr, R. Tokarz, L. Yu *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**:** 6453-6464.

Liu, Z., H. Shi, L. C. Szymczak, T. Aydin, S. Yun *et al.*, 2015 Promotion of bone morphogenetic protein signaling by tetraspanins and glycosphingolipids. PLoS Genet 11**:** e1005221.

Lowery, J. W., B. Brookshire and V. Rosen, 2016 A Survey of Strategies to Modulate the Bone Morphogenetic Protein Signaling Pathway: Current and Future Perspectives. Stem Cells Int 2016**:** 7290686.

Marchant, T. W., E. J. Johnson, L. McTeir, C. I. Johnson, A. Gow *et al.*, 2017 Canine Brachycephaly Is Associated with a Retrotransposon-Mediated Missplicing of SMOC2. Curr Biol 27**:** 1573-1584.e1576.

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

Mok, D. Z., P. W. Sternberg and T. Inoue, 2015 Morphologically defined sub-stages of C. elegans vulval development in the fourth larval stage. BMC Dev Biol 15**:** 26.

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

Morita, K., A. J. Flemming, Y. Sugihara, M. Mochii, Y. Suzuki *et al.*, 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.

585 Munoz, J., D. E. Stange, A. G. Schepers, M. van de Wetering, B. K. Koo *et al.*, 2012 The Lgr5
586 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers.
587 Embo j 31**:** 3079-3091.
588 Nonet, M. L., J. E. Staunton, M. P. Kilgard, T. Fergestad, E. Hartwieg *et al.*, 1997
589 Caenorhabditis elegans rab-3 mutant synapses exhibit impaired function and are partially
590 depleted of vesicles. J Neurosci 17**:** 8061-8073.
591 Norman, M., R. Vuilleumier, A. Springhorn, J. Gawlik and G. Pyrowolakis, 2016 Pentagone
592 internalises glypicans to fine-tune multiple signalling pathways. Elife 5.
593 Okada, I., H. Hamanoue, K. Terada, T. Tohma, A. Megarbane *et al.*, 2011 SMOC1 is essential
594 for ocular and limb development in humans and mice. Am J Hum Genet 88**:** 30-41.
595 Okkema, P. G., S. W. Harrison, V. Plunger, A. Aryana and A. Fire, 1993 Sequence requirements
596 for myosin gene expression and regulation in Caenorhabditis elegans. Genetics 135**:** 385-
597 404.
598 R Core Team, 2015 R: A language and environment for statistical computing. R
599 Foundation for Statistical Computing, pp.
600 Rainger, J., E. van Beusekom, J. K. Ramsay, L. McKie, L. Al-Gazali *et al.*, 2011 Loss of the
601 BMP antagonist, SMOC-1, causes Ophthalmo-acromelic (Waardenburg Anophthalmia)
602 syndrome in humans and mice. PLoS Genet 7**:** e1002114.
603 Ren, P., C. S. Lim, R. Johnsen, P. S. Albert, D. Pilgrim *et al.*, 1996 Control of C. elegans larval
604 development by neuronal expression of a TGF-beta homolog. Science 274**:** 1389-1391.
605 Rhiner, C., S. Gysi, E. Frohli, M. O. Hengartner and A. Hajnal, 2005 Syndecan regulates cell
606 migration and axon guidance in C. elegans. Development 132**:** 4621-4633.
607 Rocnik, E. F., P. Liu, K. Sato, K. Walsh and C. Vaziri, 2006 The novel SPARC family member
608 SMOC-2 potentiates angiogenic growth factor activity. J Biol Chem 281**:** 22855-22864.
609 Rogalski, T. M., B. D. Williams, G. P. Mullen and D. G. Moerman, 1993 Products of the unc-52
610 gene in Caenorhabditis elegans are homologous to the core protein of the mammalian
611 basement membrane heparan sulfate proteoglycan. Genes Dev 7**:** 1471-1484.
612 Rual, J. F., J. Ceron, J. Koreth, T. Hao, A. S. Nicot *et al.*, 2004 Toward improving
613 Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library.
614 Genome Res 14**:** 2162-2168.
615 Salazar, V. S., L. W. Gamer and V. Rosen, 2016 BMP signalling in skeletal development,
616 disease and repair. Nat Rev Endocrinol 12**:** 203-221.
617 Savage, C., P. Das, A. L. Finelli, S. R. Townsend, C. Y. Sun *et al.*, 1996 Caenorhabditis elegans
618 genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor
619 beta pathway components. Proc Natl Acad Sci U S A 93**:** 790-794.
620 Savage-Dunn, C., and R. W. Padgett, 2017 The TGF-beta Family in Caenorhabditis elegans.
621 Cold Spring Harb Perspect Biol 9.
622 Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair *et al.*, 2012 Fiji: an open-
623 source platform for biological-image analysis. Nat Methods 9**:** 676-682.
624 Sedlmeier, G., and J. P. Sleeman, 2017 Extracellular regulation of BMP signaling: welcome to
625 the matrix. Biochem Soc Trans 45**:** 173-181.
626 Shvab, A., G. Haase, A. Ben-Shmuel, N. Gavert, T. Brabletz *et al.*, 2016 Induction of the
627 intestinal stem cell signature gene SMOC-2 is required for L1-mediated colon cancer
628 progression. Oncogene 35**:** 549-557.

Suzuki, Y., M. D. Yandell, P. J. Roy, S. Krishna, C. Savage-Dunn *et al.*, 1999 A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in Caenorhabditis elegans. Development 126**:** 241-250.

Thawani, J. P., A. C. Wang, K. D. Than, C. Y. Lin, F. La Marca *et al.*, 2010 Bone morphogenetic proteins and cancer: review of the literature. Neurosurgery 66**:** 233-246; discussion 246.

Thomas, J. T., D. Eric Dollins, K. R. Andrykovich, T. Chu, B. G. Stultz *et al.*, 2017 SMOC can act as both an antagonist and an expander of BMP signaling. Elife 6.

Tian, C., D. Sen, H. Shi, M. L. Foehr, Y. Plavskin *et al.*, 2010 The RGM protein DRAG-1 positively regulates a BMP-like signaling pathway in Caenorhabditis elegans. Development 137**:** 2375-2384.

Tian, C., H. Shi, S. Xiong, F. Hu, W. C. Xiong *et al.*, 2013 The neogenin/DCC homolog UNC-40 promotes BMP signaling via the RGM protein DRAG-1 in C. elegans. Development 140**:** 4070-4080.

Vannahme, C., S. Gosling, M. Paulsson, P. Maurer and U. Hartmann, 2003 Characterization of SMOC-2, a modular extracellular calcium-binding protein. Biochem J 373**:** 805-814.

Vannahme, C., N. Smyth, N. Miosge, S. Gosling, C. Frie *et al.*, 2002 Characterization of SMOC-1, a novel modular calcium-binding protein in basement membranes. J Biol Chem 277**:** 37977-37986.

Vowels, J. J., and J. H. Thomas, 1992 Genetic analysis of chemosensory control of dauer formation in Caenorhabditis elegans. Genetics 130**:** 105-123.

Vuilleumier, R., A. Springhorn, L. Patterson, S. Koidl, M. Hammerschmidt *et al.*, 2010 Control of Dpp morphogen signalling by a secreted feedback regulator. Nat Cell Biol 12**:** 611-617.

Wallace, I. M., O. O'Sullivan, D. G. Higgins and C. Notredame, 2006 M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res 34**:** 1692-1699.

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

Wang, L., Z. Liu, H. Shi and J. Liu, 2017 Two Paralogous Tetraspanins TSP-12 and TSP-14 Function with the ADAM10 Metalloprotease SUP-17 to Promote BMP Signaling in Caenorhabditis elegans. PLoS Genet 13**:** e1006568.

Wang, R. N., J. Green, Z. Wang, Y. Deng, M. Qiao *et al.*, 2014 Bone Morphogenetic Protein (BMP) signaling in development and human diseases. Genes Dis 1**:** 87-105.

Wu, M., G. Chen and Y. P. Li, 2016 TGF-beta and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. Bone Res 4**:** 16009.

Yoshida, S., K. Morita, M. Mochii and N. Ueno, 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.

FIGURE LEGENDS

Fig 1. *smoc-1(0)* mutations suppress the *sma-9(0)* M lineage defect.

(**A**) Schematic representation of the BMP signaling pathway in *C. elegans*. BMP: bone morphogenetic protein. RI: type I receptor. RII: type II receptor. R-Smad: receptor-associated Smad. Co-Smad: common mediator Smad. (**B-C**) Merged DIC and GFP images of L4 stage *sma-9(cc604)* (**B**) and *smoc-1(tm7125); sma-9(cc604)* (**C**) worms carrying the *CC::gfp* coelomocyte (CC) marker. Arrows indicate M-derived CCs. Asterisks (*) denote embryonically-derived CCs.

Fig 2. SMOC-1 is conserved from *C. elegans* to human.

(**A-B**) Schematics of the *C. elegans smoc-1* gene (**A**) and the predicted SMOC-1 protein (B), respectively, showing the domain structure and the molecular lesions of various mutant alleles. SP: signal peptide. TY: thyroglobulin type I-like repeat. EC: secreted protein acidic and rich in cysteine (SPARC) extracellular calcium binding domain. (**C**) Schematic representation of *C. elegans* SMOC-1, *D. melanogaster* Pentagone, and *H. sapiens* SMOC1 and SMOC2, showing their domain structures. The two human SMOC proteins are of different lengths but share similar domain structures. FS: follistatin-like domain. (**D-E**) Alignment of the TY (**D**) and EC (**E**) domains from *C. elegans* SMOC-1, *D. melanogaster* Pentagone, and *H. sapiens* SMOC1 and SMOC2. Multiple copies of a certain domain in the same protein are numbered in order from the N-terminus to the C-terminus. Ce: *C. elegans*. Dm: *D. melanogaster*. Hs: *H. sapiens*. (**F**) Alignment of the TY domains from SMOC-1 homologs in various nematode species. Cel: *C. elegans*. Cre: *C. remanei*. Cbn: *C. brenneri*. Cbg: *C. briggsae*. Cja: *C. japonica*. Ppa: *Pristionchus pacificus*. In D-F, identical or conserved amino acids are shown on a black or grey background, respectively. Red boxes highlight residues mutated in certain *smoc-1* alleles. Blue

box indicates the residue changed in a *smoc-1* cDNA clone that rendered the protein non-functional.

Fig 3. SMOC-1 regulates body size.

(**A-C**) DIC images showing *smoc-1(tm7125)* (**A**), WT (**B**), and *smoc-1(OE)* (**C**) worms at the larval L4.3 stage. (**D**) Relative body length of developmental stage-matched WT and various *smoc-1* mutant worms. Each *smoc-1* mutant allele was outcrossed with N2 for at least three times, and two independent isolates for each allele (#s following the allele name) were used for body size measurement. The *smoc-1(+)* transgene was pMSD4*[2kb smoc-1p::smoc-1 cDNA::2kb smoc-1 3'UTR]*. The body length of WT worms was set to 1.0. Error bars represent 95% confidence interval (CI). An ANOVA followed by TukeyHSD was used to test for differences between genotypes. *** $P<0.0001$.

Fig 4. SMOC-1 functions through the BMP ligand to positively regulate BMP signaling.

(**A-B**) Relative body length of developmental stage-matched WT and various mutant worms, including double mutants between *smoc-1(tm7125)* and null mutants in the BMP pathway (**A**), and double mutants between *smoc-1(OE)* and null mutants in the BMP pathway (**B**). Two independent isolates for each double mutant combination were used for body size measurement. The body length of WT worms was set to 1.0. Error bars represent 95% CI. (**C**) Representative GFP images showing RAD-SMAD reporter expression in hypodermal nuclei of WT, *lon-2(e678)*, and *smoc-1(OE)* worms, respectively. (**D**) Boxplot showing the relative RAD-SMAD GFP fluorescence intensity in WT (set to 1.0), *lon-2(e678)*, and two independent isolates of *smoc-1(OE)* worms. Each data point represents an average of the GFP fluorescence intensity

from five hypodermal nuclei in one worm. Approximately 40 worms were examined per genotype. For panels A, B and D, an ANOVA followed by TukeyHSD was used to test for differences between genotypes. ND: no difference. *** $P<0.0001$.

Fig 5. SMOC-1 antagonizes LON-2/glypican in regulating body size.

Relative body length of developmental stage-matched WT (set to 1.0) and various mutant worms, including double mutants between *smoc-1(tm7125)* null and *lon-2(e678)* null (**A**), and double mutants between *smoc-1(OE)* and *lon-2(e678)* null (**B**). The body size of *smoc-1(tm7125)*; *lon-2(e678)* double mutants is similar to that of WT animals, while *smoc-1(OE); lon-2(e678)* double mutants are longer than either one. Error bars represent 95% CI. An ANOVA followed by TukeyHSD was used to test for differences between genotypes. ND: no difference. * $P<0.01$. ** $P<0.001$. *** $P<0.0001$.

Fig 6. *smoc-1* is expressed in multiple tissues and its intestinal expression is positively regulated by BMP signaling.

(**A**) Expression of *smoc-1* cDNA under different regulatory elements to test for rescue of the body size phenotype of *smoc-1(tm7125)* worms. For each construct, two independent transgenic lines were examined and the data were combined and averaged. Body sizes are relative to *smoc-1(tm7125)* mutant worms (set to 1.0), and all measurements were done on the same day. Error bars represent 95% CI. *** $P<0.0001$. (**B-F**) Merged GFP and DIC images of wildtype worms (**B-E**) and a *sma-6(jj1)* mutant worm (**F**) carrying the integrated *smoc-1* transcriptional reporter *jjIs4688* (Table 1). GFP expression is detectable a bean stage embryo (**E**), and in cells of the pharynx (**B-C**), intestine (**B**), and posterior hypodermis (**B, D**) in a WT larva. GFP expression in

the intestine, but not in the pharynx or posterior hypodermis, is significantly reduced in *sma-6(jj1)* (**F**). Images are side views with anterior to the left and dorsal up. (**G**) Quantification of the penetrance of L4 stage animals showing intestinal expression of the *smoc-1* transcriptional reporter in wildtype and various BMP pathway mutants. Two independent isolates were assessed for each gene tested.

Fig 7. *smoc-1* functions cell non-autonomously to regulate body size and M lineage development.

Tissue specific expression of *smoc-1* cDNA to test for rescue of the body size (**A**) or Susm (**B**) phenotype of *smoc-1(tm7125)* worms. *smoc-1* cDNA was driven by each specific promoter to allow expression in a given tissue. All constructs used the *unc-54* 3'UTR. For each construct, two independent transgenic lines were examined and the measurements were averaged. (**A**) Body sizes are relative to *smoc-1(tm7125)* mutant worms (set to 1.0), and all measurements were done on the same day. Error bars represent 95% CI. (**B**) The Susm phenotype was scored in the background of *smoc-1(tm7125); sma-9(cc604); CC::gfp*. Error bars represent standard error. ** $P<0.001$. *** $P<0.0001$. ND: no difference.

Fig 8. Human SMOC proteins can partially rescue the Susm phenotype of *smoc-1(0)* mutants.

(**A**) Schematics of SMOC homologs tested for function in *C. elegans*. Solid black outline indicates *C. elegans* protein sequences. Dashed grey line indicates human protein sequences. All ORFs were cloned into the same vector with the same regulatory elements (2kb *smoc-1* promoter and *unc-54* 3'UTR), and each construct was tested for rescue of Susm (**B**) and body size (**C**)

760 phenotype of *smoc-1(tm7125)* mutants. Two independent lines were assayed for each construct.
761 (**B**) The Susm phenotype was scored in the background of *smoc-1(tm7125); sma-9(cc604);*
762 *CC::gfp*. Error bars represent standard error. (**C**) Body sizes are relative to WT worms (set to
763 1.0), and all measurements were done on the same day. Error bars represent 95% CI. * $P<0.01$.
764 ** $P<0.001$. *** $P<0.0001$. ND: no difference.
765
766 **Fig 9. A model for SMOC-1 function in the BMP pathway.**
767 SMOC-1 acts through the BMP ligand DBL-1/BMP, and in part by antagonizing LON-
768 2/glypican, to promote BMP signaling. BMP signaling in turn promotes the intestinal expression
769 of *smoc-1,* thus creating a positive feedback loop.

Table 1. Strains used in this study.

Strain ID	Genotype
Original *sma-9* suppressor strains from the EMS screen	
LW0040	*arIs37[secreted CC::gfp] I; cup-5(ar465) III; sma-9(cc604) X*
LW2697	*arIs37[secreted CC::gfp] I; cup-5(ar465) III; smoc-1(jj65) V; sma-9(cc604) X*
LW2732	*arIs37[secreted CC::gfp] I; cup-5(ar465) III; smoc-1(jj85) V; sma-9(cc604) X*
LW2731	*arIs37[secreted CC::gfp] I; sma-4(jj70) cup-5(ar465) III; smoc-1(jj180) V; sma-9(cc604) X*
LW3874	*arIs37[secreted CC::gfp] I; cup-5(ar465) III; smoc-1(jj109) V; sma-9(cc604) X*
LW3927	*arIs37[secreted CC::gfp] I; cup-5(ar465) III; smoc-1(jj115) V; sma-9(cc604) X*
LW3906	*arIs37[secreted CC::gfp] I; cup-5(ar465) III; smoc-1(jj139) V; sma-9(cc604) X*
Strains with different *smoc-1* alleles	
LW4477	*smoc-1(tm7000) V [6x outcrossed, isolate 3.23]*
LW4478	*smoc-1(tm7000) V [6x outcrossed, isolate 4.5]*
LW4479	*smoc-1(tm7125) V [6x outcrossed, isolate 5.2]*
LW4480	*smoc-1(tm7125) V [6x outcrossed, isolate 7.24]*
LW4766	*smoc-1(jj65) V [5x outcrossed, isolate 2.13]*
LW4487	*smoc-1(jj85) V [3x outcrossed, isolate 1.13]*
LW4555	*smoc-1(jj180) V [5x outcrossed, isolate 5.4]*
LW4556	*smoc-1(jj180) V [5x outcrossed, isolate 5.8]*
LW5623	*smoc-1(jj115) V [3x outcrossed, isolate 3.5]*
LW5624	*smoc-1(jj115) V [3x outcrossed, isolate 13.3]*
LW5129	*jjIs5119[pMSD4.4(smoc-1p::smoc-1 cDNA:: smoc-1 3'UTR)+pCFJ90(myo-2p::mCherry)] I 3x outcrossed, isolate 1.3, also known as smoc-1(OE)*
LW5130	*jjIs5119[pMSD4.4(smoc-1p::smoc-1 cDNA:: smoc-1 3'UTR)+pCFJ90(myo-2p::mCherry)] I 3x outcrossed, isolate 2.5, also known as smoc-1(OE)*
Strains for examining the M lineage phenotypes of *smoc-1* mutants	
LW0081	*ccIs4438 [intrinsic CC:::gfp] III; ayIs2[egl-15p::gfp] IV; ayIs6[hlh-8p::gfp] X*
LW4420	*ccIs4438[intrinsic CC::gfp] III; ayIs2[egl-15p::gfp] IV; smoc-1(tm7000) V; ayIs6[hlh-8p::gfp] X*
LW4422	*ccIs4438[intrinsic CC::gfp] III; ayIs2[egl-15p::gfp] IV; smoc-1(tm7125) V; ayIs6[hlh-8p::gfp] X*
LW4442	*arIs37[secreted CC::gfp] I; ccIs4438[intrinsic CC::gfp] III; ayIs2[egl-15p::gfp] IV; smoc-1(tm7125) V; sma-9(cc604) ayIs6[hlh-8p::gfp] X*
LW4443	*arIs37[secreted CC::gfp] I; ccIs4438[intrinsic CC::gfp] III; ayIs2[egl-15p::gfp] IV; smoc-1(tm7000) V; sma-9(cc604) ayIs6[hlh-8p::gfp] X*
LW4457	*arIs37[secreted CC::gfp] I; smoc-1(jj180) V; sma-9(cc604) X*

LW4834	*arIs37[secreted CC::gfp] I; ccIs4438[intrinsic CC::gfp] III; smoc-1(tm7125) V; sma-9(cc604) ayIs6[hlh-8p::gfp] X*
Strains carrying RAD-SMAD reporter	
LW2433	*jjIs2433[pCXT51(5*RLR::deleted pes-10p::gfp) + LiuFD61(mec-7p::rfp)] X, isolate 1, also known as RAD-SMAD reporter*
LW3467	*dbl-1(wk70) V; jjIs2433[RAD-SMAD] X*
LW3468	*lon-2(e678) jjIs2433[RAD-SMAD] X*
LW5604	*jjIs5119[smoc-1(OE)] I; jjIs2433[RAD-SMAD] X, isolate 1*
LW5605	*jjIs5119[smoc-1(OE)] I; jjIs2433[RAD-SMAD] X, isolate 2*
Strains for body size measurement	
LW1856	*sma-6(jj1) II*
LW5498	*sma-3(tm4625) III [4x outcrossed, isolate 8.3]*
LW5499	*sma-3(tm4625) III [4x outcrossed, isolate 8.6]*
LW3346	*sma-3(jj3) III*
LW4774	*dbl-1(ok3749) V*
LW3471	*lon-2(e678) X*
LW4703	*sma-6(jj1) II; smoc-1(tm7125) V, isolate 9.3.1*
LW4704	*sma-6(jj1) II; smoc-1(tm7125) V, isolate 9.6.1*
LW4590	*sma-3(jj3) III; smoc-1(tm7125) V, isolate 13.12*
LW4595	*sma-3(jj3) III; smoc-1(tm7125) V, isolate 5.9.5*
LW5344	*dbl-1(ok3749) smoc-1(tm7125) V, isolate 1.1.1*
LW5345	*dbl-1(ok3749) smoc-1(tm7125) V, isolate 4.7.8*
LW4617	*smoc-1(tm7125) V; lon-2(e678) X, isolate 6.10.3.6*
LW4618	*smoc-1(tm7125) V; lon-2(e678) X, isolate 6.7.3.7*
LW5241	*jjIs5119[smoc-1(OE)] I; dbl-1(ok3749) V, isolate 2.17*
LW5263	*jjIs5119[smoc-1(OE)] I; dbl-1(ok3749) V, isolate 3.11*
LW5621	*jjIs5119[smoc-1(OE)] I; sma-3(tm4625) III, isolate 1.2*
LW5622	*jjIs5119[smoc-1(OE)] I; sma-3(tm4625) III, isolate 2.3*
LW5294	*jjIs5119[smoc-1(OE)] I; lon-2(e678) X, isolate 5*
LW5295	*jjIs5119[smoc-1(OE)] I; lon-2(e678) X, isolate 6*
Strains for assaying the dauer phenotype	
DR40	*daf-1(m40) IV*
DR609	*daf-1(m213) IV*
CB1372	*daf-7(e1372) III*
LW5288	*daf-1(m40) IV; smoc-1(tm7125) V isolate 7.11B*
LW5289	*daf-1(m40) IV; smoc-1(tm7125) V isolate 16.16B*

LW5286 *daf-1(m213) IV; smoc-1(tm7125) V isolate 2.4*

LW5287 *daf-1(m213) IV; smoc-1(tm7125) V isolate 2.8*

LW5290 *daf-7(e1372) III; smoc-1(tm7125) V isolate 6.12*

LW5306 *daf-7(e1372) III; smoc-1(tm7125) V isolate 6.6*

LW5291 *daf-1(m213) IV; lon-2(e678) X isolate 2.3*

LW5292 *daf-1(m213) IV; lon-2(e678) X isolate 2.4*

LW5293 *daf-7(e1372) III; lon-2(e678) X isolate 6*

LW5285 *daf-7(e1372) III; lon-2(e678) X isolate 15*

Strains carrying the *smoc-1* reporter constructs

LW4688 *jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV, 3x outcrossed, isolate 13.1*

LW4694 *jjIs4694[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] V, 3x outcrossed, isolate 19.1*

LW4764 *sma-3(jj3) III; jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV, isolate 3.2*

LW4765 *sma-3(jj3) III; jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV, isolate 9.3*

LW4724 *sma-6(jj1) II; jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV, isolate 11.4.2*

LW4725 *sma-6(jj1) II; jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV, isolate 11.7.2*

LW4728 *jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV; lon-2(e678), isolate 1.3.1*

LW4729 *jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV; lon-2(e678), isolate 2.1.2*

LW5520 *jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV; smoc-1(tm7125) V, isolate 1.7*

LW5521 *jjIs4688[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] I or IV; smoc-1(tm7125) V, isolate 2.5*

LW4878 *jjIs3900[pJKL1066.3(hlh-8p::nls::mCherry::lacZ)+ pCFJ90(myo-2p::mCherry)] IV; jjIs4694[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] V, isolate 1.2*

LW4879 *jjIs3900[pJKL1066.3(hlh-8p::nls::mCherry::lacZ)+ pCFJ90(myo-2p::mCherry)] IV; jjIs4694[pJKL1139.2(smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4] V, isolate 2.4*

LW5656 *jjEx5656[pJKL1201(5kb smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4]*

LW5657 *jjEx5657[pJKL1201(5kb smoc-1p::4xnls::gfp::unc-54 3'UTR)+pRF4]*

LW5658 *jjEx5658[pJKL1202(5kb smoc-1p::4xnls::gfp::2kb smoc-1 3'UTR)+pRF4]*

LW5659 *jjEx5658[pJKL1202(5kb smoc-1p::4xnls::gfp::2kb smoc-1 3'UTR)+pRF4]*

Table 2. Plasmid constructs generated in this study.

Plasmid name	Construct information	
Translational and transcriptional reporter constructs		
pJKL1128	*2kb smoc-1p::smoc-1 genomic::2kb smoc-1 3'UTR*	
pMSD4	*2kb smoc-1p::smoc-1 cDNA::2kb smoc-1 3'UTR*	
pJKL1138	*2kb smoc-1p::smoc-1 cDNA::unc-54 3'UTR*	
pJKL1139	*2kb smoc-1p::4xnls::gfp::unc-54 3'UTR*	
pJKL1201	*5kb smoc-1p::4xnls::gfp::unc-54 3'UTR*	
pJKL1202	*5kb smoc-1p::4xnls::gfp::2kb smoc-1 3'UTR*	
Tissue specific expression constructs		
pJKL1137	*hlh-8p::smoc-1 cDNA::unc-54 3'UTR*	M lineage
pJKL1136	*hlh-8p::smoc-1 cDNA-S103P::unc-54 3'UTR*	M lineage
pMSD6	*elt-3p::smoc-1 cDNA::unc-54 3'UTR*	hypodermis
pMSD7	*myo-2p::smoc-1 cDNA::unc-54 3'UTR*	pharyngeal muscles
pMSD8	*myo-3p::smoc-1 cDNA::unc-54 3'UTR*	body wall muscles
pMSD9	*rab-3p::smoc-1 cDNA::unc-54 3'UTR*	pan neurons
pMSD18	*ifb-2p::smoc-1 cDNA::unc-54 3'UTR*	intestine
Constructs to express human *SMOC* genes		
pJKL1150	*2kb smoc-1p::huSMOC1 ORF::smoc-1 3'UTR*	
pJKL1151	*2kb smoc-1p::huSMOC2 ORF::smoc-1 3'UTR*	
pJKL1178	*2kb smoc-1p::CelSP::huSMOC1 chimera::unc-54 3'UTR*	
pJKL1179	*2kb smoc-1p::CelSP::huSMOC2 chimera::unc-54 3'UTR*	

Table 3. Mutations in *smoc-1* suppress the *sma-9(0)* M lineage defects.

Genotype	Susm penetrance[a] (# of animals examined)
sma-9(cc604)	--
smoc-1(jj65); sma-9(cc604)	84% (N=255)[b]
smoc-1(jj85); sma-9(cc604)	78% (N=240)[b]
smoc-1(jj180); sma-4(jj70); sma-9(cc604)	98% (N=80)[b,c]
smoc-1(jj180); sma-9(cc604)	98% (N=319)[c]
sma-4(jj70); sma-9(cc604)	0% (N>100)[c]
sma-4(e729); sma-9(cc604)	100% (N=61)[d]
smoc-1(tm7000); sma-9(cc604)	97% (N=134)
smoc-1(tm7125); sma-9(cc604)	98% (N=686)
smoc-1(tm7000)/jj65 or +/jj65; sma-9(cc604)	67% (N=51)[e]
smoc-1(tm7000)/jj85 or +/jj85; sma-9(cc604)	46% (N=24)[e]
smoc-1(tm7000)/jj180 or +/jj180; sma-9(cc604)	58% (N=26)[e]
smoc-1(jj109); sma-9(cc604)	99% (N=107)
smoc-1(jj115); sma-9(cc604)	100% (N=95)
smoc-1(jj139); sma-9(cc604)	100% (N=128)
smoc-1(tm7125); sma-9(cc604); jjEx4490[smoc-1p::smoc-1 genomic::smoc-1 3'UTR], line 1	32% (N=111)
smoc-1(tm7125); sma-9(cc604); jjEx4491[smoc-1p::smoc-1 genomic::smoc-1 3'UTR], line 2	26% (N=101)
smoc-1(tm7125); sma-9(cc604); jjEx4810[smoc-1p::smoc-1 cDNA::smoc-1 3'UTR], line 1	2% (N=278)
smoc-1(tm7125); sma-9(cc604); jjEx4811[smoc-1p::smoc-1 cDNA::smoc-1 3'UTR], line 2	1% (N=498)
smoc-1(tm7125); sma-9(cc604); jjEx4812[smoc-1p::smoc-1 cDNA::unc-54 3'UTR], line 1	17% (N=481)
smoc-1(tm7125); sma-9(cc604); jjEx4813[smoc-1p::smoc-1 cDNA::unc-54 3'UTR], line 2	23% (N=792)
smoc-1(tm7125); sma-9(cc604); jjEx4650[hlh-8p::smoc-1 cDNA::unc-54 3'UTR], line 1	15% (N=186)
smoc-1(tm7125); sma-9(cc604); jjEx4612[hlh-8p::smoc-1 cDNA::unc-54 3'UTR], line 3	17% (N=214)
smoc-1(tm7125); sma-9(cc604); jjEx4620[hlh-8p::smoc-1 cDNA-S103P::unc-54 3'UTR], line 1	86% (N=95)
smoc-1(tm7125); sma-9(cc604); jjEx4674[hlh-8p::smoc-1 cDNA-S103P::unc-54 3'UTR], line 2	92% (N=100)

[a] The Susm penetrance refers to the percent of animals with 1-2 M-derived CCs as scored by the *CC::GFP* reporter.
[b] Data taken from (LIU *et al.* 2015).
[c] The *jj70* strain described in our previous publication (LIU *et al.* 2015) carries a mutation in *sma-4(*S110L), as well as a mutation in *smoc-1*(Q180Stop). To avoid confusion, we have designated the *sma-4* mutation as *jj70*, and the mutation in *smoc-1* as *jj180*. As shown here, *sma-4(jj70)* failed to suppress *sma-9(0),* while *smoc-1(jj180)* suppressed *sma-9(0)*.
[d] Data taken from (FOEHR *et al.* 2006).
[e] Complementation tests were performed by crossing *tm7000/+; cc604* males with *jj65 (jj85 or jj180); cc604* hermaphrodites and scoring the cross progeny for the number of CCs. All progeny would have 4 CCs if the tested alleles complemented each other, while ~50% of the progeny would have 6 CCs if the tested alleles failed to complement each other. The partial dominance of each of the *jj* alleles tested (LIU *et al.* 2015) may have contributed to the observed percentage being slightly above 50%.

Table 4. SMOC-1 does not play a significant role in the TGFβ dauer pathway.

Genotype	15°C % Daf-c (n)	20°C % Daf-c (n)	25°C % Daf-c (n)
smoc-1(tm7125)	0 (858)	0 (828)	0 (863)
jjIs5119[smoc-1(OE)]	0 (541)	0 (792)	0 (574)
daf-7(e1372)	30.2±4.5 (348)	92.8±2.9 (794)	99.8±0.3 (954)
daf-7(e1372); smoc-1(tm7125) #1	33.8±14.8 (142)	82.5±9.3 (748)*	99.6±0.3 (1102)
daf-7(e1372); smoc-1(tm7125) #2	26.4±7.7 (148)	72.0±8.3 (343)*	100 (1115)
daf-1(m40)	0 (485)	44.9±7.4 (1059)	100 (964)
daf-1(m40); smoc-1(tm7125) #1	0 (589)	57.2±16.4 (1567)*	99.9±0.2 (970)
daf-1(m40); smoc-1(tm7125) #2	0 (483)	24.0±8.8 (721)*	100 (518)
daf-1(m213)	0 (469)	99.4±0.6 (867)	100 (1174)
daf-1(m213); smoc-1(tm7125) #1	0 (603)	98.2±3.1 (649)	100 (719)
daf-1(m213); smoc-1(tm7125) #2	0 (544)	99.4±0.5 (676)	100 (378)

n: number of worms scored at each temperature; from a total of 5 plates per genotype assayed at each condition. For each double mutant combination, two independent isolates (#1 and #2) were examined.

% Daf-c: mean dauer formation percentage ± standard deviation.

*: $p<0.05$, as calculated by an ANOVA and TukeyHSD, between double mutant and the corresponding *daf* single mutant at the specified temperature.

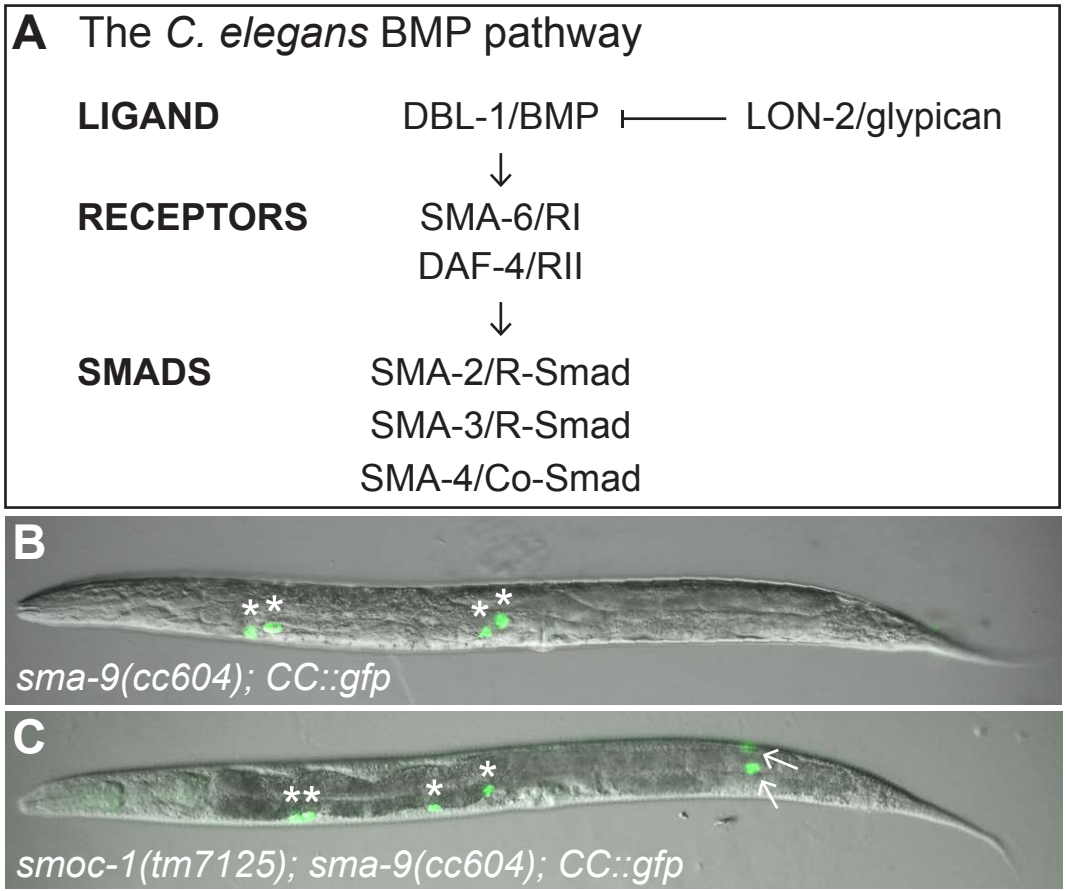
Table 5. LON-2 does not play a significant role in the TGFβ dauer pathway.

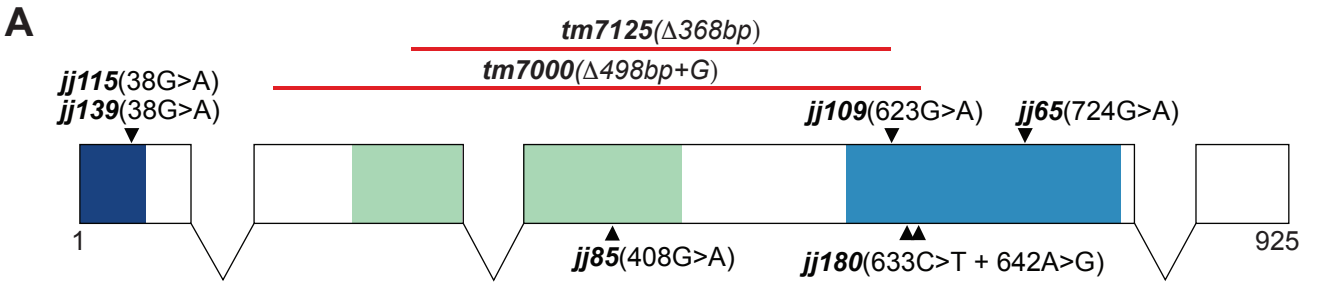
Genotype	15°C % Daf-c (n)	20°C % Daf-c (n)	25°C % Daf-c (n)
lon-2(e678)	0 (807)	0 (799)	0 (1090)
daf-7(e1372)	22.7±32.1 (141)	83.3±6.5 (257)	100 (357)
daf-7(e1372); lon-2(e678)) #1	25.3±8.7 (435)	59.4±16.4 (239)*	100 (749)
daf-7(e1372); lon-2(e678) #2	31.7±4.8 (249)	66.7±11.7 (426)	100 (840)
daf-1(m213)	0.3±1.0 (313)	97.9±26.0 (570)	100 (853)
daf-1(m213); lon-2(e678) #1	0.2±0.3 (575)	92.1±2.9 (643)	100 (1149)
daf-1(m213); lon-2(e678) #2	0.5±0.8 (654)	77.7±5.0 (515)	100 (832)

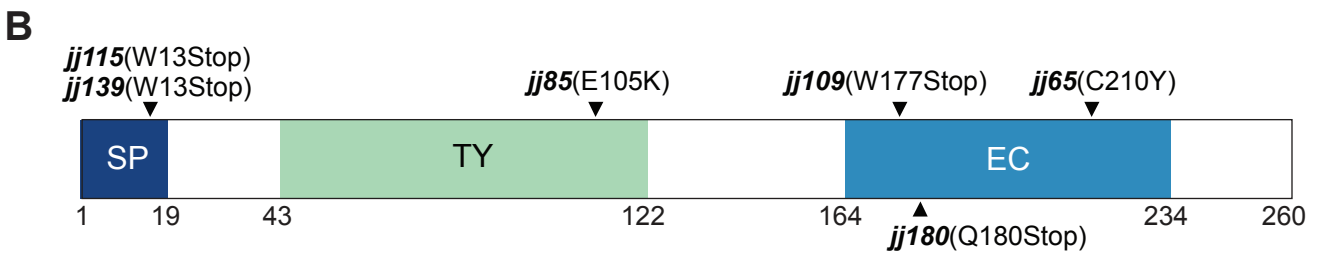
n: number of worms scored at each temperature; from a total of 5 plates per genotype assayed at each condition. For each double mutant combination, two independent isolates (#1 and #2) were examined.

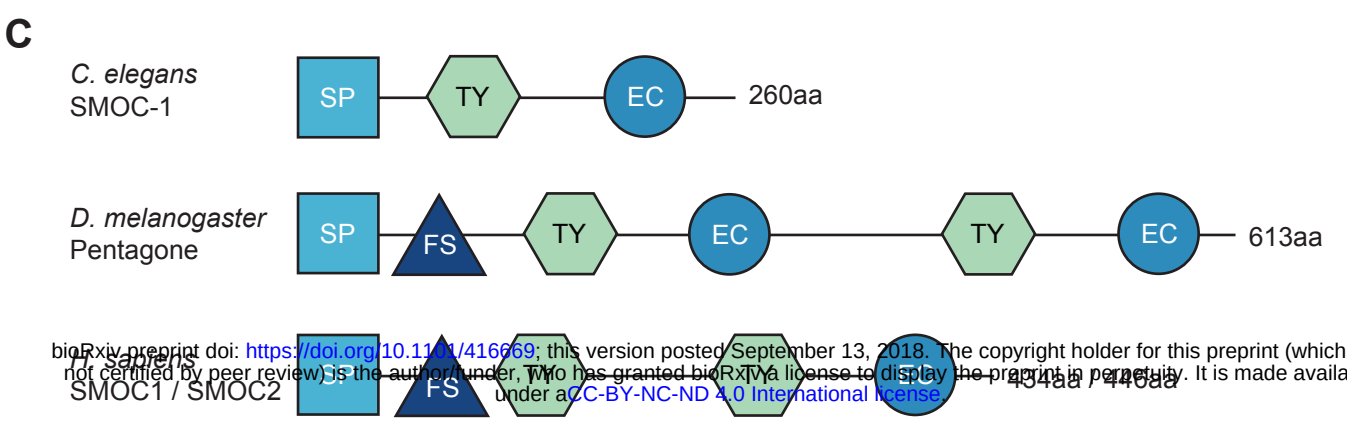
% Daf-c: mean dauer formation percentage ± standard deviation.

*: $p<0.05$, as calculated by an ANOVA and TukeyHSD, between double mutant and the corresponding *daf* single mutant at the specified temperature.


A
The C. elegans BMP pathway
LIGAND
DBL-1/BMP
LON-2/glypican
RECEPTORS
SMA-6/RI
DAF-4/RII
SMADS
SMA-2/R-Smad
SMA-3/R-Smad
SMA-4/Co-Smad
B
sma-9(cc604); CC::gfp
C
smoc-1(tm7125); sma-9(cc604); CC::gfp







D TY domain alignment

```
Ce_SMOC1_TY      53 CASQRRKALKRKTD-GDARIYIPTCSPKNSLLYDKVQCYDVS--IYCWCV
Dm_Pent_TY1      80 CLEAVKFARRQQER--DPGYFVPRCRKD-GN-FAAMQCYGNN--G-CWCS
Dm_Pent_TY2     396 CWMDQSVTLEEQGHGGKSVLFVPQCLPD-GR-YQRIQCYSSTSTSYCWCV
Hs_SMOC1_TY1     95 CRLERAQALEQAKK-PQEAVFVPECGED-GS-FTQVQCHTYT--GYCWCV
Hs_SMOC1_TY2    227 CDQERQSALEEAQQNPREGIVIPECAPG-GL-YKPVQCHQST--GYCWCV
Hs_SMOC2_TY1     90 CVAERKYTQEQARK-EFQQVFIPECNDD-GT-YSQVQCHSYT--GYCWCV
Hs_SMOC2_TY2    216 CDQEHQSALEEAKQPKNDNVVIPECAHG-GL-YKPVQCHPST--GYCWCV
                    *  ..  ... . .     ...*.*  . .  .  .**.  .  ..***.

Ce_SMOC1_TY     100 DELSGEPKLGSSTTRG--------------KPKCE 120
Dm_Pent_TY1     123 DSQ-GRPIADDNKQFRRKGKLRCRANRRDRRRLAS 156
Dm_Pent_TY2     444 NEDTGKSIPGTSVKNK--------------RPQCD 464
Hs_SMOC1_TY1    140 TPD-GKPISGSSVQNK--------------TPVCS 159
Hs_SMOC1_TY2    273 LVDTGRPLPGTSTRYV--------------MPSCE 293
Hs_SMOC2_TY1    135 TPN-GRPISGTAVAHK--------------TPRCP 154
Hs_SMOC2_TY2    262 LVDTGRPIPGTSTRYE--------------QPKCD 282
                    .*.. ...   .                . ..
```

S103P *jj85*(E105K)

E EC domain alignment

```
Ce_SMOC1_EC     133 RRNNRCKEKKRTRFLRRLVSTLKSEMIMSGINAT----------KV---
Dm_Pent_EC1     177 TAHRTCSKSDRSQFNTNLMRMFRNEA-QSFFRQP----------SL---
Dm_Pent_EC2     470 RPMKGCTEPRKTQFLKELKAYLNTSLLPSSTTGS----------NSSMW
Hs_SMOC1_EC     310 RELPGCPEGKKMEFITSLLDALTTDMVQAINSAAPTGGGRFSEPDPS--
Hs_SMOC2_EC     299 RQLQGCPGAKKHEFLTSVLDALSTDMVHAASDPSS-SSGRLSEPDPS--
                    . . .*  . ... *.. ..  . .... .    ..           .
```

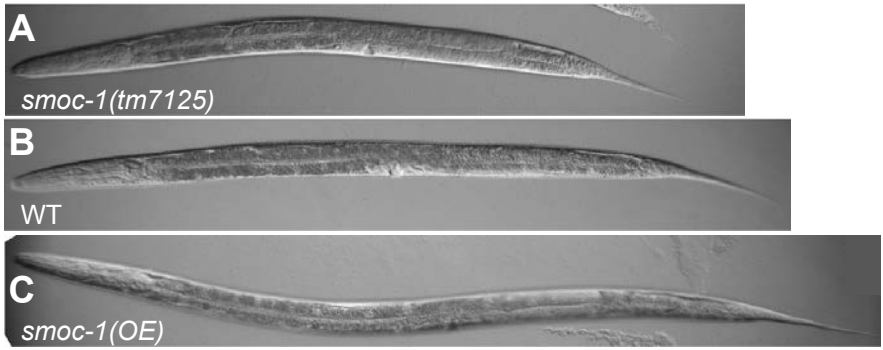
jj65(C210Y)

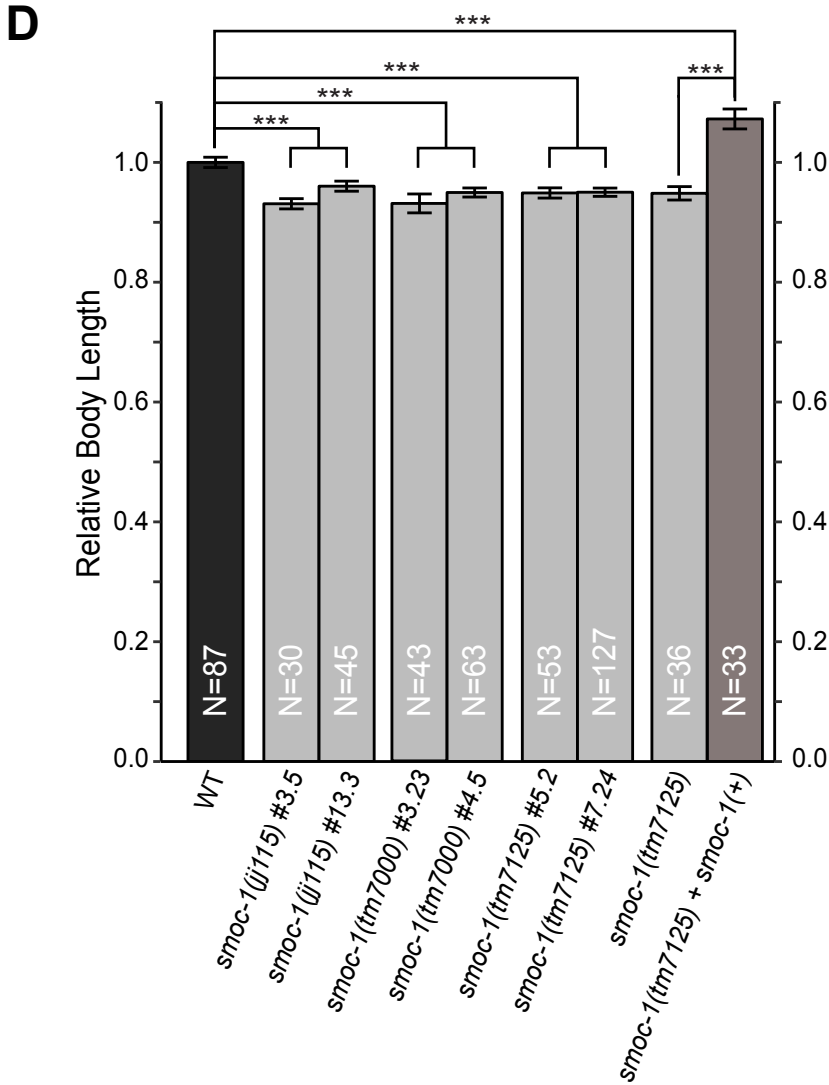
```
Ce_SMOC1_EC     169 -SRDSAIRWKFNQLNINHNNVLERSEWKPFKSVLLEWKNVRQCSRNLFK
Dm_Pent_EC1     212 -SDSHILEWQFSKLDTNGNKLLDRQEIRELKKVLRRNVKPRRCGRTFGK
Dm_Pent_EC2     510 TDDERIATLSFVYLDKNKNKSWDRREWKNFRDLVTSASHLRRCGKKMPR
Hs_SMOC1_EC     358 TLEERVVHWYFSQLDSNSSNDINKREMKPFKRYVKKKAKPKKCARRFTD
Hs_SMOC2_EC     346 TLEERVVHWYFKLLDKNSSGDIGKKEIKPFKRFLRKKSKPKKCVKKFVE
                    . ........ *  *. * .  ....*..........  ....*.... .

Ce_SMOC1_EC     218 CDLNKDRKLTFDEWRKCIVQEINRVPAK 245
Dm_Pent_EC1     261 CDVTKDANLNWLEWSVCFTKEFHNRSAV 288
Dm_Pent_EC2     260 CDVNGDKKISLAEWLNCL-QATPRESAT 586
Hs_SMOC1_EC     408 CDLNKDKVISLPELKGCLGVSKEGRL-V 434
Hs_SMOC2_EC     396 CDVNNDKSISVQELMGCLGVAKEDGKAD 423
                    **...*. ... *.  *.         .
```

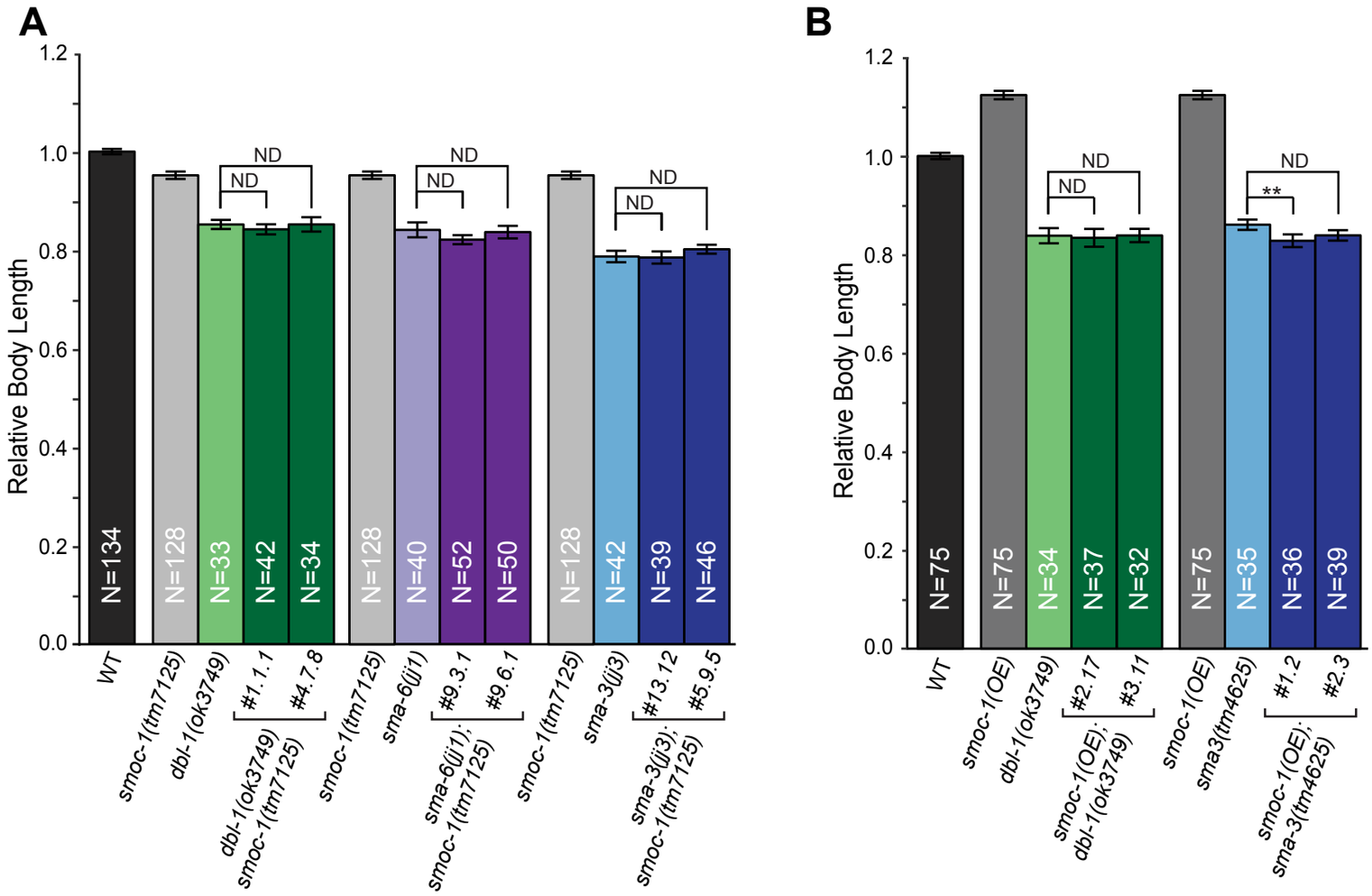
F Nematode TY domain alignment

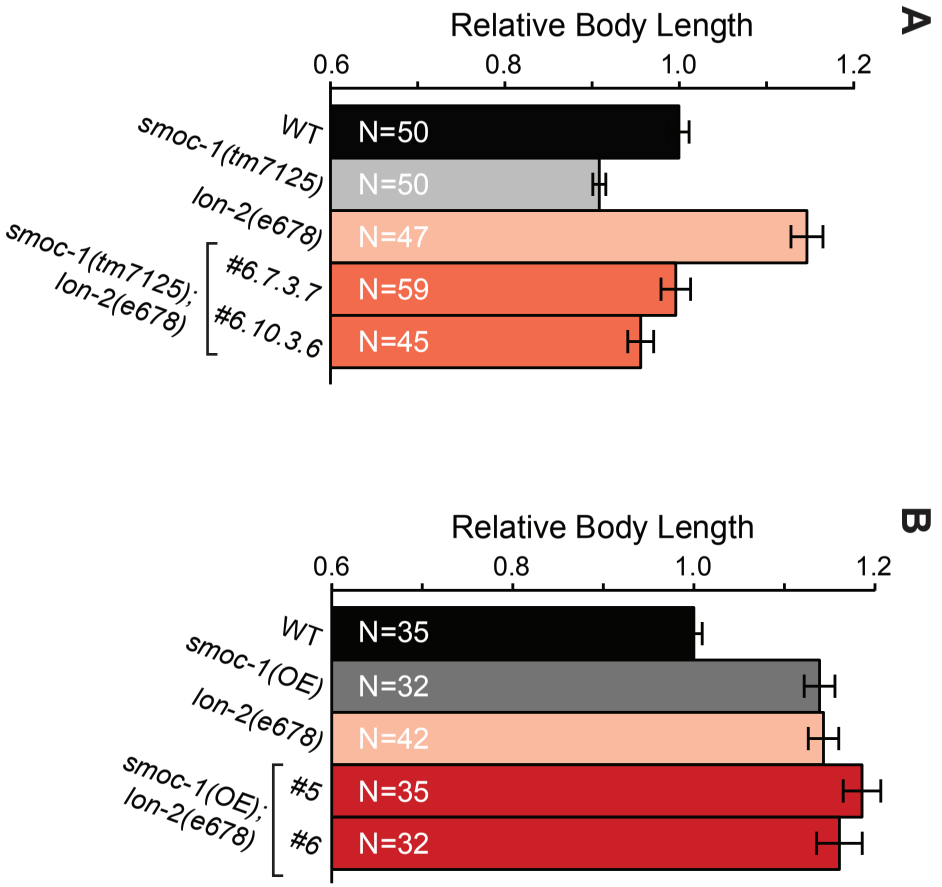
```
Ce_SMOC1          1 CASQRRKALKRKTDGDA--RIYIPTCSPKNSLLYDKVQCYDVSIYCWCVD
Cre_26999         1 CASQRRKALKRKTDGDA--RIYIPTCSPKNSLLYDKVQCYEVSAYCWCVD
Cbn_20462         1 CASQRRKALKRKTDGDS--RIYIPTCSPKNSLLYDKIQCYDVSAYCWCVD
Cbg_23276         1 CASQRRKALKRKTNGDS--RIYIPTCSPKNALLYDKVQCYDVSAYCWCVD
Cja_07338         1 CASQQRKALNRKNAGDS--KIYVPTCSAKNSLLYDKVQCYDMSAYCWCVD
Ppa_34808         1 CEQARSDLLKQMEGRQSSSVAYLPQCDMRDESLYRRLQCHGKEV-CWCVD
                    *.......*........   ..*.*.*.....**...**.....*****

Ce_SMOC1         49 ELSGEPKLGSSTTRGKPKCE 68
Cre_26999        49 ELSGEPKIGSSTSRGKPKCE 68
Cbn_20462        49 ELSGEPKIGSSTSRGKPKCE 68
Cbg_23276        49 ELSGEPKIGSSTSRGRPKCE 68
Cja_07338        49 EISGEPKIGSSTSRGKPKCQ 68
Ppa_34808        50 IVTGDPVTD--------ECE 61
                    ..*.*.*...........*.
```

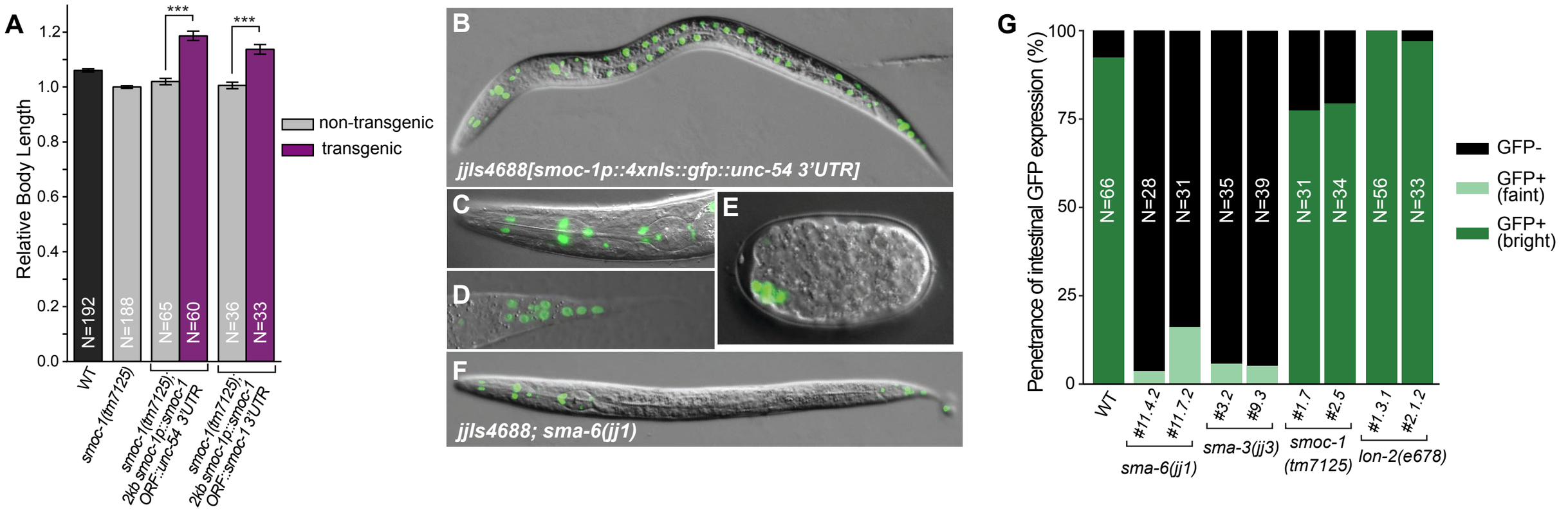
S103P *jj85*(E105K)

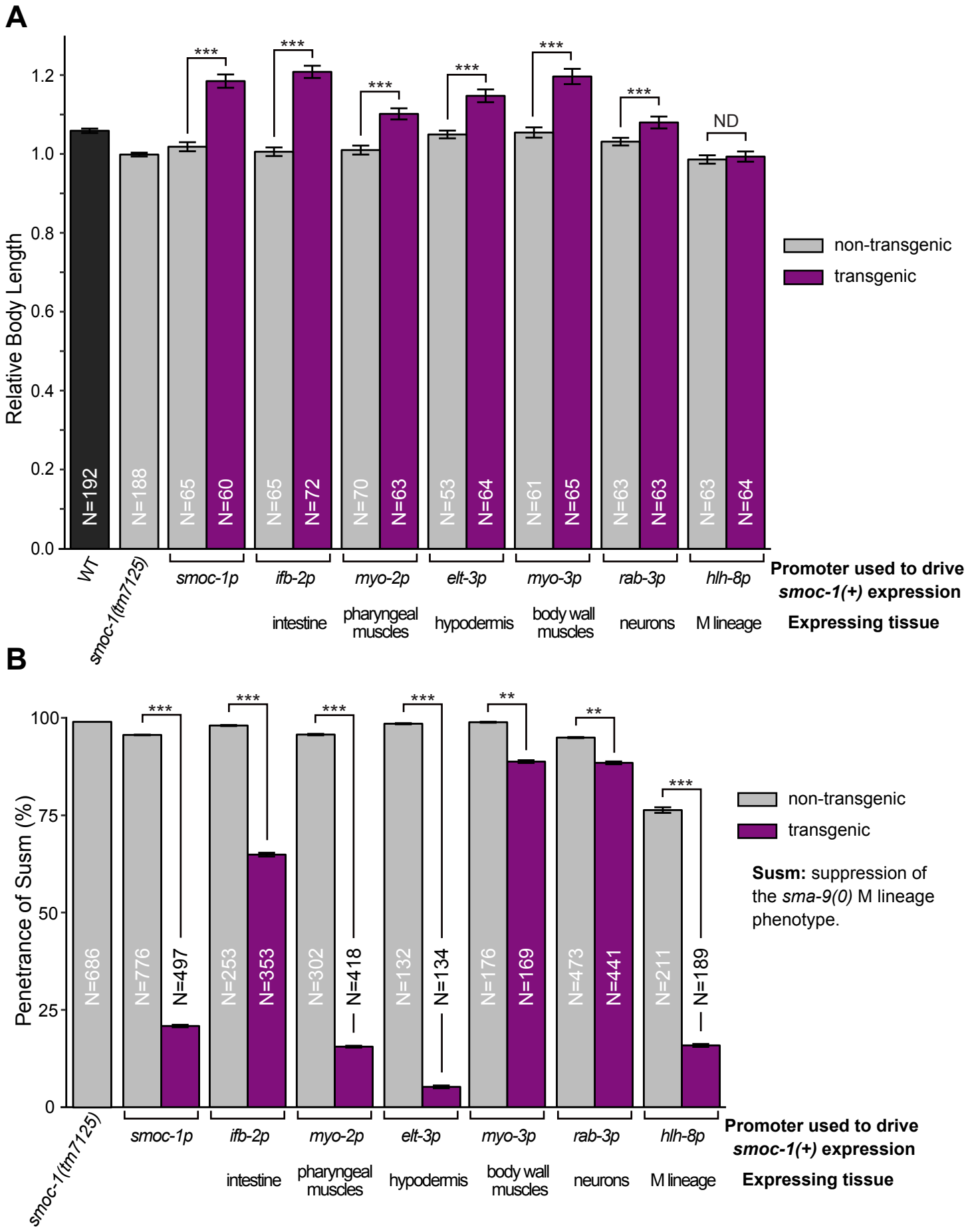

A
smoc-1(tm7125)
B
WT
C
smoc-1(OE)

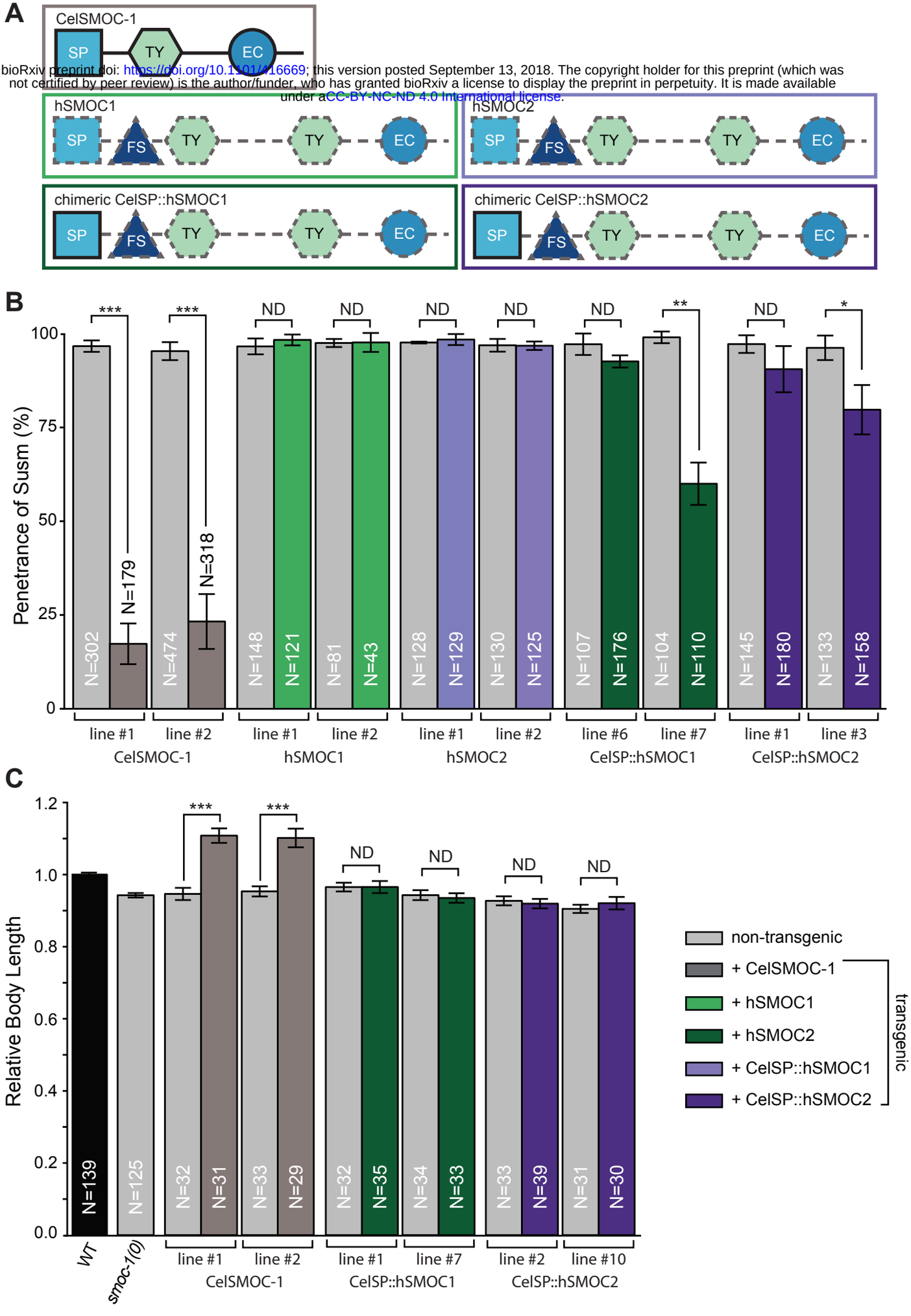

D
***
***
***
***
***
Relative Body Length
0.0
0.2
0.4
0.6
0.8
1.0
N=87
N=30
N=45
N=43
N=63
N=53
N=127
N=36
N=33
WT
smoc-1(jj115) #3.5
smoc-1(jj115) #13.3
smoc-1(tm7000) #3.23
smoc-1(tm7000) #4.5
smoc-1(tm7125) #5.2
smoc-1(tm7125) #7.24
smoc-1(tm7125)
smoc-1(tm7125) + smoc-1(+)




A
Relative Body Length
0.6
0.8
1.0
1.2
WT
N=50
smoc-1(tm7125)
N=50
lon-2(e678)
N=47
smoc-1(tm7125);
lon-2(e678)
#6.7.3.7
N=59
#6.10.3.6
N=45
B
Relative Body Length
0.6
0.8
1.0
1.2
WT
N=35
smoc-1(OE)
N=32
lon-2(e678)
N=42
smoc-1(OE);
lon-2(e678)
#5
N=35
#6
N=32


A
Relative Body Length
1.2
1.0
0.8
0.6
0.4
0.2
0.0
***
***
N=192
N=188
N=65
N=60
N=36
N=33
non-transgenic
transgenic
WT
smoc-1(tm7125)
smoc-1(tm7125); 2kb smoc-1p::smoc-1 ORF::unc-54 3'UTR
smoc-1(tm7125); 2kb smoc-1p::smoc-1 ORF::smoc-1 3'UTR
B
jjIs4688[smoc-1p::4xnls::gfp::unc-54 3'UTR]
C
D
E
F
jjIs4688; sma-6(jj1)
G
Penetrance of intestinal GFP expression (%)
100
75
50
25
0
N=66
N=28
N=31
N=35
N=39
N=31
N=34
N=56
N=33
GFP-
GFP+ (faint)
GFP+ (bright)
WT
#11.4.2
#11.7.2
sma-6(jj1)
#3.2
#9.3
sma-3(jj3)
#1.7
#2.5
smoc-1 (tm7125)
#1.3.1
#2.1.2
lon-2(e678)


A
Relative Body Length
1.2
1.0
0.8
0.6
0.4
0.2
0.0
***
ND
N=192
N=188
N=65
N=60
N=65
N=72
N=70
N=63
N=53
N=64
N=61
N=65
N=63
N=63
N=63
N=64
non-transgenic
transgenic
WT
smoc-1(tm7125)
smoc-1p
ifb-2p
myo-2p
elt-3p
myo-3p
rab-3p
hlh-8p
Promoter used to drive smoc-1(+) expression
intestine
pharyngeal muscles
hypodermis
body wall muscles
neurons
M lineage
Expressing tissue
B
Penetrance of Susm (%)
100
75
50
25
0
**
N=686
N=776
N=497
N=253
N=353
N=302
N=418
N=132
N=134
N=176
N=169
N=473
N=441
N=211
N=189
non-transgenic
transgenic
Susm: suppression of the sma-9(0) M lineage phenotype.
smoc-1(tm7125)
smoc-1p
ifb-2p
myo-2p
elt-3p
myo-3p
rab-3p
hlh-8p
Promoter used to drive smoc-1(+) expression
intestine
pharyngeal muscles
hypodermis
body wall muscles
neurons
M lineage
Expressing tissue



A model for SMOC-1 function in the BMP pathway