Development 133, 1023-1033 (2006) doi:10.1242/dev.02275

RNA-binding proteins SOP-2 and SOR-1 form a novel PcG-like complex in *C. elegans*

Tingting Zhang*, Yinyan Sun*, E Tian*, Hansong Deng, Yuxia Zhang, Xin Luo, Qingchun Cai, Huayi Wang, Jijie Chai and Hong Zhang†

We describe the identification and characterization of a novel *PcG* gene in *C. elegans*, *sor-1*, which is involved in global repression of Hox genes. *sor-1* encodes a novel protein with an RNA-binding activity. We provide evidence that SOR-1 and the previously identified RNA-binding protein SOP-2 may constitute an RNA-binding complex in Hox gene repression. SOR-1 and SOP-2 directly interact with each other and are colocalized in nuclear bodies. The localization of SOR-1 depends on SOP-2. Surprisingly, homologs of SOR-1 and SOP-2 are not found in other organisms, including the congeneric species *C. briggsae*, suggesting an unexpected lack of evolutionary constraint on an essential global gene regulatory system.

KEY WORDS: *sor-1*, *PcG*, Nuclear bodies, SAM domain, *C. elegans*

INTRODUCTION

Epigenetic gene silencing involves the formation of heritable repressive chromatin structures that are established in a stepwise manner, through initiation, histone modification and propagation along the chromatin fiber. Establishment of such chromatin structure appears to involve the concerted action of RNA and interacting proteins (Heard, 2004; Lippman and Martienssen, 2004). The RNA components can be a non-coding structural RNA, such as Xist RNA, which functions in mammalian X chromosome inactivation (Pannings and Jaenisch, 1998), or a small interfering RNA (siRNA) derived from the RNAi machinery, which is employed in the silencing of genes within the heterochromatic domains in *S. pombe* and *Drosophila* (Volpe et al., 2002; Pal-Bhadra et al., 2004). These RNA components could be responsible for targeting the interacting protein complex to specific loci and/or in later maintaining the repressed state of the genes (Akhtar, 2003; Lippman and Martienssen, 2004).

The epigenetic silencing of Hox genes mediated by Polycomb group (PcG) proteins is also thought to involve the formation of localized repressive chromatin structures (Levine et al., 2004; Lund and van Lohuizen, 2004). Hox genes encode homeodomain-containing transcription factors that specify the positional identities of cells along the anteroposterior axis during multicellular organism development (Gellon and McGinnis, 1998). In *PcG* mutants, Hox genes are ectopically expressed in regions outside their normal expression domains, resulting in homeotic transformations, in which the body structures are duplicated or lost. Two conserved PcG complexes have been identified, the ESC/E(Z) complex, containing Extra sex combs and Enhancer of zeste, and the PRC1 complex, the core components of which include Polycomb (PC), Posterior sex combs (PSC), Polyhomeotic (PH) and RING1 (Levine et al., 2004). The ESC/E(Z) complex has been shown to associate with histone deacetylases and also contains an intrinsic histone methyltransferase activity that specifically methylates H3 at lysine 9 (K9) and lysine 27 (K27) (van der Vlag and Otte, 1999; Muller et al., 2002). The PRC1 complex functions as an E3 ubiquitin ligase that is specific for H2A (de Napoles et al., 2004; Wang et al., 2004). Recent observations suggest that, like other forms of epigenetic gene silencing, PcG-mediated Hox gene repression might also involve RNA (Sun and Zhang, 2004). For example, the ESC/E(Z) complex and the PRC1 complex are required for X chromosome inactivation in mammals and they are recruited to the X chromosome in a Xist structural RNA-dependent manner, although their direct interaction has not been established (Silva et al., 2003; Plath et al., 2003; Plath et al., 2004).

PcG-mediated Hox gene repression is an ancient mechanism, conserved from flies to mammals. However, the nematode *C. elegans* apparently lacks PRC1 complex genes. Furthermore, although the ESC/E(Z) complex is found (e.g. *mes-2* and *mes-6*), mutations in these genes result in only mild homeotic defects (Korf et al., 1998; Holdeman et al., 1998; Ross and Zarkower, 2003). Nevertheless, *C. elegans* Hox genes are subject to global repression. This repression is mediated by a novel gene, *sop-2*. In *sop-2* mutants, the onset of Hox gene expression is normal but is subsequently expressed ectopically in diverse body regions. Like the components of the PRC1 complex, SOP-2 forms nuclear bodies, called SOP-2 bodies (Zhang et al., 2003). The formation of SOP-2 bodies is tightly correlated with its function and its formation appears to require the RNA-binding activity of SOP-2 (Zhang et al., 2003; Zhang et al., 2004a; Zhang et al., 2004b), although the nature of the RNA under physiological conditions remains to be elucidated (Zhang et al., 2004b).

Identification of additional components in the SOP-2 nuclear bodies will provide insight into the role of SOP-2 in Hox gene repression. Although SOP-2 bears little sequence similarity to PRC1 complex proteins, components of the PRC1 complex are also localized into distinct nuclear bodies, called PcG bodies, and also have RNA-binding activity (Netter et al., 2001; Zhang et al., 2004b). The study of the SOP-2/RNA complex is likely to shed light on the salient features of PcG-mediated gene repression. In this study, we describe the identification and characterization of a *PcG*-like gene in *C. elegans*, *sor-1*, that encodes a novel RNA-binding protein with limited regions of similarity to the mouse PcG protein Rae28. SOR-1 and SOP-2 colocalize in SOP-2 nuclear bodies and direct interact with each other. Our studies reveal that SOR-1 and SOP-2 define a putative PcG-like complex in global repression of Hox genes in *C. elegans*. Remarkably, neither SOR-1 nor SOP-2 is conserved in

National Institute of Biological Sciences, Beijing, People's Republic of China.

*These authors contributed equally to this work
†Author for correspondence (e-mail: zhanghong@nibs.ac.cn)

Accepted 5 January 2006

other organisms, not even in the congeneric species *C. briggsae*, suggesting a surprising lack of evolutionary constraint on an ancient regulatory system.

MATERIALS AND METHODS

Strains

Most strains carrying the *him-5(e1490)* mutation, which produces a high frequency of males in self progenies. The following strains were used in this work.

LGII: *sop-2(bx91)*, *muIs16(mab-5::gfp, dpy-20)*
LGIII: *sor-1(bp1)*; *sor-1(bp2)*; *sor-1(bp3)*; *mab-5(e1239) egl-5(n945)*
LGV: *bxIs14(pkd-2::gfp, pha-1(+))*
LGX: *bxIs13(egl-5::gfp*, *lin-15(+))*

The *bx91* mutation is temperature sensitive. At 15°C, animals have no obvious defects. At 25°C, animals arrest at L1 and L2 stages. Adult *sop-2(bx91)* animals were shifted from 15°C to 25°C and the arrested early larvae were examined for the expression pattern of Hox genes.

Isolation, mapping and cloning of *sor-1*

sor-1 mutations were isolated in screens for mutants with ectopic expression of *mab-5::gfp* and *egl-5::gfp* reporters. Approximately 17,000 haploid genomes were screened. *bp1*, *bp2* and *bp3* are located at the same genetic locus by genetic mapping and non-complementation experiments. Three factor mapping placed *sor-1* between *unc-32(0.00)* and *sma-3(–0.93)*, about –0.23, on LGIII. Six out of eight Sma nonUnc and seven out of 28 Unc nonSma recombinants carried *sor-1* mutation. PCR products from ZK1236.3 were injected into + *sor-1(bp2)* +/*unc-32* + *sma-3* together with transformation marker pRF4*(rol-6(su1006))*, and its ability of rescuing the early larval lethality of *sor-1(bp2)* was assessed. The heterozygotes transformants gave rise to adult *sor-1(bp2)* males and hermaphrodites, while 100% of *sor-1(bp2)* animals without the transgene arrested at L1 and L2 stages.

cDNA clones of *sor-1,* yk526e9 and yk336g5, were kindly provided by Dr Kohara (NIG, Japan) and were sequenced to confirm the predicted exon/intron junctions. The 5′ end of the *sor-1* cDNA was obtained by RT-PCR using SL1 sense primer and an antisense primer that is specific to *sor-1*. SL1 is a trans-spliced leader that is present at the 5′ end of many *C. elegans* cDNAs. The *sor-1* mutations identified in our genetic screens were determined by sequencing the corresponding sequence of the *sor-1* locus.

RNA interference

Single-stranded RNA (ssRNA) was transcribed from the T7 and T3-flanked PCR templates (ZK1236 nucleotides 6709-7498) with MEGAscript T3 and T7 kits (Ambion). The ssRNAs were then annealed, and injected into *muIs16*, *bxIs13*, *bxIs14*, *mab-5 egl-5; bxIs14* and *sop-2(bx91); bxIs13* animals. F1 progenies generated 4 hours after injection were scored for larval lethality, ectopic expression of Hox genes, or generation of anterior rays, as shown by the expression of *pkd-2::gfp* reporter.

sor-1::gfp reporter gene

The *sor-1::gfp* reporter was constructed by PCR fusion based approach (Hobert, 2002). The fused PCR products were derived from two overlapping PCR DNA fragments. One contained the DNA derived from ZK1236 (nucleotides 3272 to 9530), which includes a 3 kb promoter region and the entire ORF of *sor-1*. Another one contained the *gfp* and the *unc-54* 3′UTR from pPD95.67. The PCR products were co-injected with pBX-1*(pha-1+)* into *pha-1* mutant worms and the transformants were analyzed.

Preparation of antibody to SOR-1 protein

The *sor-1* cDNA corresponding to the N terminus of SOR-1 (amino acid 230 to 390) was cloned into the pET28 expression vector. This his-tagged fusion protein produced by *E. coli* BL21 was purified to be used as an immunogen in rabbits. The antisera were first absorbed with bacterial acetone powder, followed by NAB protein A spin purification kit (Pierce).

Indirect immunofluorescence

Embryos were obtained from well-fed adult hermaphrodites. The permeabilization of embryos and young larvae was performed by Freeze-Cracking methods (Albertson, 1984). The freeze-cracked slides were fixed, blocked and incubated with anti-SOR-1 antibody at a final dilution of 1:400 at room temperature for 2-4 hours. The worms were then washed three times and incubated with Rhodamine Red-X-conjugated goat anti-rabbit IgG. The specificity of SOR-1 antibody was demonstrated by lack of staining in control experiments with pre-immune sera, with antisera that were preincubated with 0.25 mg of purified SOR-1 fusion protein, or in *sor-1(RNAi)* embryos. *sop-2(bx91)* L2 larvae were shifted from 20°C to 25°C, and the embryos derived form these animals were used for immunostaining.

GST pull-down experiments

Constructs encoding GST-SOR-1 and GST-SOP-2 were made by subcloning portions of the cDNAs of SOR-1 and SOP-2 into pGEX-4T-1. Fusion proteins were expressed in BL21 cells and purified with glutathione sepharose resin according to the manufacturer's recommendation (Pharmacia). For in vitro systems of ^{35}S-labelled proteins, the corresponding *sor-1* and *sop-2* were cloned into pcDNA3 as templates for transcription and translation (TNT Coupled Reticulocyte Lysate System, Promega). The GST fusion proteins (200 ng) were incubated with ^{35}S-labelled protein and 10 μl glutathione sepharose beads in binding buffer [25 mM Tris.Cl (pH 7.6), 150 mM NaCl, 1 mM DTT, 0.5% Triton X-100, 10% Glycerol, 1 mM PMSF] for 2 hours at 4°C. The reactions were then washed for four times with 1 ml binding buffer. Bound proteins were analyzed by SDS-PAGE and analyzed by autoradiograph.

For in vivo GST pull-down assay, SOP-2 (residues 58-140) and SOR-1 (48-200) were cloned into the vectors of pGEX-6P-1 and PET30a, respectively, and co-expressed in *E. coli* strain BL21(DE3). GST-4B resin was first used for purification of the complex. GST was removed by thrombin and then the digested complex was subjected to ion exchange column (Amersham-Pharmacia) for further purification.

Gel filtration assay

The complex was purified to more than 95% homogeneity in assay buffer (10 mM Tris, pH 8.0, 3 mM DTT, 100 mM NaCl). The complex (1.0 ml, about 2.0 mg complex) was subjected to gel filtration analysis (Superdex200, Amersham-Pharmacia). Samples from the peak corresponding to the complex were visualized by SDS-PADE and stained with Coomassie Blue.

EMSA and RNA-binding assays

RNA-binding reactions contained 20 mM HEPES (pH 7.6), 100 mM KCl, 2 mM EDTA, 0.01% NP40, 1 mM DTT, labeled RNA fragment (20,000 cpm). Ten times cold irrelevant ssRNA, dsRNA and 100 times yeast tRNA (mass excess) were used for competition experiments. Reactions (25 μl) were incubated on ice for 20 minutes and electrophoresed on 4% native TBE PAGE gel and analyzed by autoradiograph.

RESULTS

sor-1 acts as a *PcG*-like gene in global repression of Hox genes in *C. elegans*

To identify components involved in maintaining the integrity of expression domains of Hox genes in *C. elegans*, strains expressing the Hox gene reporter, *egl-5::gfp* or *mab-5::gfp*, were used to screen for mutants with expanded expression domains of *egl-5* and *mab-5*, respectively. Three alleles of a new genetic locus, termed *sor-1(sop-2 related-1) (bp1*, *bp2*, *bp3)*, were isolated from the total 17,000 mutagenized genomes. In early wild-type larvae, *egl-5* expression is restricted to cells in the tail region (Fig. 1A) (Ferreira et al., 1999). In *sor-1* mutants, *egl-5* is ectopically expressed in head region, and sometimes in the anterior seam cells (Fig. 1B). Similarly, the expression domain of *mab-5* is greatly expanded. During early wild-type larval development, expression of *mab-5* is limited to the cells of posterior body region, including P9-P12, V5, and V6 (Fig. 1C) (Kenyon et al., 1997). In *sor-1* mutants, *mab-5::gfp* is ectopically expressed in the head, and in the syncytial hypodermal cell 7 (hpy7) (Fig. 1D). Taken together, these observations suggest that wild-type *sor-1* is required for global repression of Hox genes.

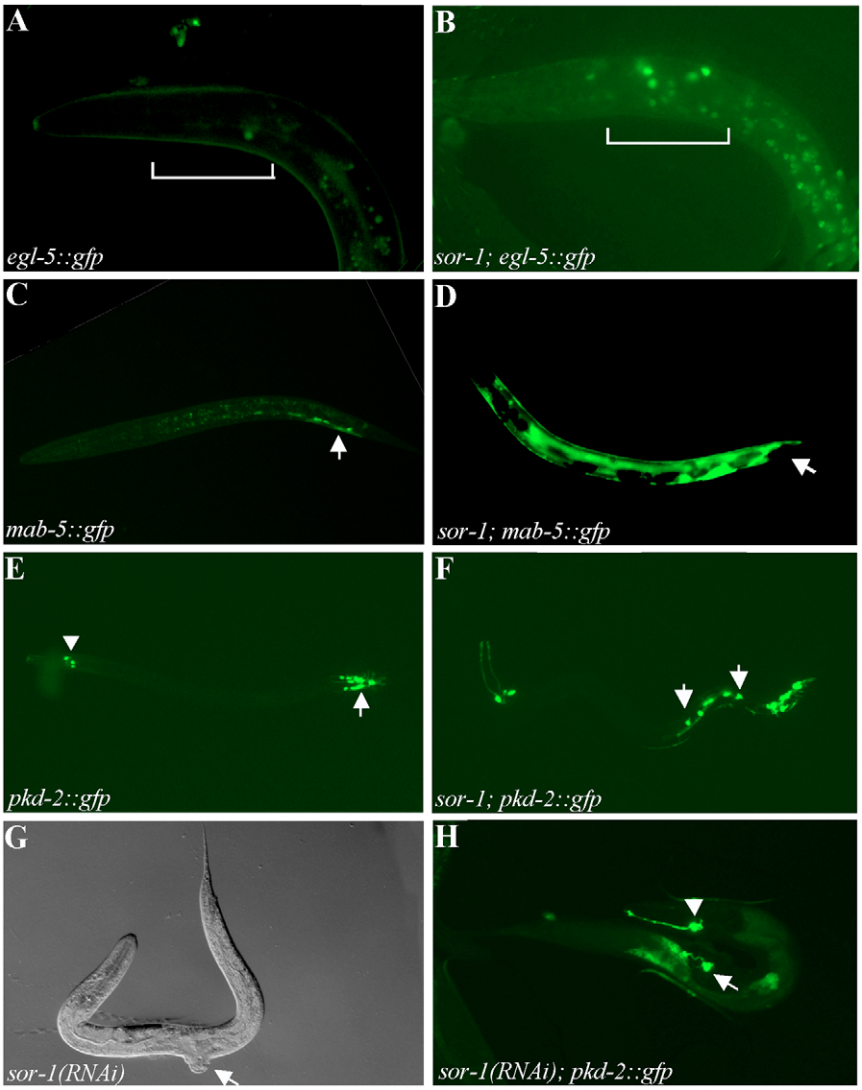

Fig. 1. Ectopic expression of Hox genes and other defects in *sor-1* loss-of-function mutants. (**A**) Expression of *egl-5* in a wild-type L2 larva. The expression of *egl-5* is restricted to the tail (not shown) and is absent in the head (bracket). (**B**) Ectopic expression of *egl-5* in a *sor-1* mutant. *egl-5* is ectopically expressed in many head neurons (marked with bar). (**C**) Expression of *mab-5* in a wild-type larva. The expression of *mab-5* is confined to the posterior region (arrow). (**D**) Ectopic expression of *mab-5* in a *sor-1* mutant. *mab-5* is expressed in the head neurons and hypodermal cells. The expression of *mab-5::gfp* is absent from the tail region (arrow). (**E**) Normal expression of *pkd-2::gfp* in all nine pairs of B-type ray neurons (arrow) and in four head neurons (arrowhead) in a wild-type male. PKD-2::GFP marks both the cell body and the axon of neurons. (**F**) Ectopic expression of *pkd-2::gfp* in a *sor-1(RNAi)* male, indicating that ectopic rays are generated in the anterior body region. Five ectopic rays are located between the arrows. (**G**) *sor-1* mutant hermaphrodites have a protruding vulva phenotype (arrow). (**H**) Partial hermaphrodite-to-male sexual transformation. An ectopic male ray (arrow), as visualized with *pkd-2::gfp*, is generated in a *sor-1* hermaphrodite. This animal also contains a *pkd-2::gfp*-positive neuron in the head (arrowhead), which is normally expressed in four male-specific head neurons.

The ectopic expression of Hox genes in *sor-1* mutants could result either from ectopic Hox gene initiation or from later failure in maintaining the repressed state of Hox genes. To distinguish between these possibilities, we determined the onset of ectopic expression of *egl-5::gfp* and *mab-5::gfp* in *sor-1* mutant embryos. As in the developing larvae, *egl-5::gfp* and *mab-5::gfp* reporters are expressed in a limited number of cells. Inappropriate expression domains of *egl-5* and *mab-5* are first detected when embryogenesis reached the threefold stage, when most cell lineages are completed. At this time, *egl-5::gfp* begins to be expressed in the head region and *mab-5::gfp* is ectopically expressed in hyp7. Hence, as in the *PcG* mutants of other organisms, the expression of Hox genes in *sor-1* mutants is correctly initiated but later becomes derepressed.

sor-1 mutations cause anterior to posterior homeotic transformations

To determine whether the endogenous Hox genes are ectopically expressed in *sor-1* mutants (like their reporter genes), we examined whether homeotic transformations, phenotypic readouts of mis-expression of Hox genes, occur in *sor-1* mutants. We studied the development of a row of epidermal seam cells in males, in which the role of Hox genes *mab-5* and *egl-5* in cell fate specification is well characterized (Emmons, 1999). During wild-type male development, the three most-posterior seam cells, V5, V6 and T, give rise to nine sensory rays. *mab-5* is expressed in the lineages of V5 and V6, and is required for the generation of rays from these cells (Salser and Kenyon, 1996). *egl-5* is expressed in the posterior branches of the V6 lineage where it specifies the identities of the V6 rays (Ferreira et al., 1999). The anterior seam cells, from V1 to V4, which do not express *mab-5* and *egl-5*, generate three parallel longitudinal cuticular ridges, known as alae. Ectopic expression or loss of function of Hox genes *mab-5* and *egl-5* results in homeotic transformation of the fates adopted by seam cells V1 to V6. The early larval lethality of *sor-1* mutants identified in our screens prevents us from analyzing the role of *sor-1* in male ray development, an event that occurs later in larval development. We therefore took advantage of *sor-1(RNAi)*, which at the appropriate concentration causes weaker loss-of-function defects in the animals developed from the eggs laid in the first few hours after injection. *sor-1* dsRNAs (gene identification see below) were injected into worms carrying a *pkd-2::gfp* reporter, which is specifically expressed in one of the two ray neurons and serves as a ray marker (Fig. 1E) (Barr and Sternberg, 2000). We found that in *sor-1(RNAi)* males, *pkd-2::gfp* is ectopically expressed in the anterior body region instead of being confined to the posterior (89%, n=18) (Fig. 1F), indicating that ectopic rays are generated from anterior seam

cells. By contrast, ectopic anterior rays are completely absent in *sor-1(RNAi) mab-5 egl-5* males (*n*=13). Thus, loss of function of *sor-1* causes anterior to posterior cell fate transformation in a Hox gene-dependant manner.

sor-1 mutants exhibit pleiotropic defects

In addition to ectopic expression of Hox genes, *sor-1* mutant animals display many other defects. All *sor-1* mutants identified in our genetic screens arrest as homozygotes at L1 and L2 larval stages. In *sor-1(RNAi)* hermaphrodites, vulva defects are apparent, such as bursting vulva and protruding vulva (Fig. 1G). In addition, 20% of *sor-1(RNAi)* mutants (*n*=41) show partial hermaphrodite-to-male sexual transformation, revealed by inappropriate expression of *pkd-2::gfp*, a marker of male-specific neurons fates (Fig. 1H). *sor-1(RNAi)* hermaphrodites are also sterile because of defects in gonads and germline development. These defects cannot be readily attributed to misregulated expression of Hox genes and suggest that *sor-1* may have other targets in addition to Hox genes.

sor-1 and *sop-2* act synergistically in Hox gene repression

The defects of *sor-1* mutants are similar to those previously described for *sop-2* mutants (Zhang et al., 2003). Both *sor-1* and *sop-2* are involved in global repression of Hox genes and in regulating the expression of non-homeotic genes. To assess the interactions between *sop-2* and *sor-1*, we generated *sop-2; sor-1* double mutants by performing *sor-1(RNAi)* in *sop-2(bx91)* animals. *sor-1(RNAi)* is likely to recapitulate a null or strong loss-of-function phenotype in the animals developed from the eggs laid in 4-48 hours after injection, as the same extent of ectopic expression of Hox genes was observed in *sor-1 (RNAi)* and *sor-1 (bp1)* (in which about two-thirds of SOR-1 is deleted) mutant animals. Shifitng *sop-2(bx91ts)* to non-permissive temperature (25°C) at L4 stage appears to cause null defects in the next generation, as the same extent of defects were observed as those in *sop-2(bp7)* animals, a putative null allele of *sop-2*, in which a stop codon mutation occurs at position 80, leading to a deletion of the 655 C-terminal residues of SOP-2. We chose RNAi because both *sop-2(bp7)* and *sor-1(bp1)* homozygous animals arrest at L1 to L2 larval stages and can be maintained only as heterozygous. We examined the genetic interactions between *sor-1* and *sop-2* by comparing the detailed expression pattern of *egl-5* in *sor-1(RNAi); sop-2(bx91)* double mutants with that observed in *sor-1(RNAi)* and *sop-2(bx91)* single mutant animals. First, we determined the onset of ectopic expression of Hox genes. In *sor-1* and *sop-2* single mutants, Hox genes are not ectopically expressed until the developing embryos reach the threefold stage. By contrast, ectopic expression of *egl-5* in the head region was observed earlier, at the 'pretzel' stage in *sop-2(bx91); sor-1(RNAi)* embryos (*n*=15) (Fig. 2E,F), during which neither *sor-1(RNAi)* (*n*=6) nor *sop-2(bx91)* (*n*=3) shows ectopic expression of *egl-5* (Fig. 2A-D). Thus, *sor-1* and *sop-2* function synergistically in maintaining the repressed state of Hox genes in early embryogenesis.

Next, we examined the spatial limits of ectopic expression domains of Hox genes in *sop-2; sor-1* animals. In *sor-1* and *sop-2* single mutants, *egl-5* is ectopically expressed in head neurons and in tail cells. The number of cells expressing the *egl-5* reporter, however, is dramatically increased in *sop-2(bx91); sor-1(RNAi)* double mutants. In *sop-2(bx91); sor-1(RNAi)* mutants, the average number of cells expressing *egl-5* is 47 in the head and 28 in the tail, compared with 13 in the head and eight in the tail in *sor-1(RNAi)* mutants, and with 32 in the head and 12 in the tail in *sop-2* mutants. The increased number of *egl-5*-positive cells in *sop-2; sor-1(RNAi)* mutants is not an additive effect of *sor-1* and *sop-2*, as the expression domains in the head and tail are dramatically expanded (Table 1; Fig. 2G-I). Moreover, in 89% (*n*=19) of the *sop-2(bx91); sor-1(RNAi)* animals, *egl-5* is seen in the mid-body region, including the ventral cord and seam cells (Fig. 2J), whereas it is observed in only 13% (*n*=39) of *sor-1* mutants and 24% (*n*=37) of *sop-2* mutants. Therefore, *sor-1* and *sop-2* act synergistically in some body regions to repress Hox gene expression.

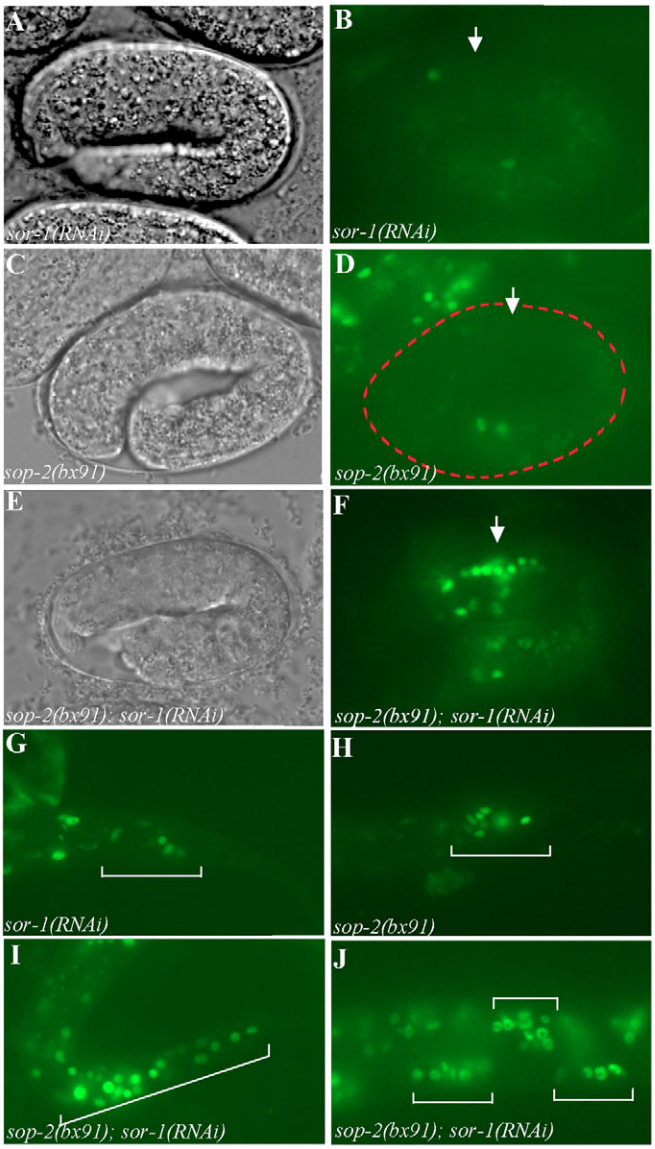

Fig. 2. Synergistic effects of *sor-1* and *sop-2*. (A-F) An earlier onset of ectopic expression of *egl-5* in *sop-2; sor-1* mutants. Lack of expression of *egl-5::gfp* in the head region in a 'pretzel' stage embryo (arrow) in *sor-1* (**A**,**B**) and *sop-2* (**C**,**D**) mutants. Ectopic expression of *egl-5::gfp* in the head region in a 'pretzel' stage *sop-2; sor-1(RNAi)* embryo (**E**,**F**) (arrow). *egl-5* is expressed in the tail region in all mutant embryos. Nomarski images (A,C,E); Fluorescence images (B,D,F). (**G**) Expression of *egl-5::gfp* in the tail of a *sor-1(RNAi)* mutant (bracket). (**H**) Expression of *egl-5::gfp* in the tail of a *sop-2* mutant (bracket). (**I**) Dramatically expanded expression domain of *egl-5::gfp* in the tail of a *sop-2; sor-1(RNAi)* mutant (bracket). (**J**) Ectopic expression of *egl-5::gfp* in the middle body region in a *sop-2; sor-1(RNAi)* mutant (marked with a bar). Irregular fluorescence particles are gut autofluorescence.

Table 1. Genetic interactions between the SOP-2/SOR-1 complex and the MES-2/MES-6 complex

	Number of cells expressing *egl-5::gfp*		Lethality (%)					
	Head	Tail	L1	L2	L3	L4	Adult	*n*
sor-1(RNAi)	12.5/*n*=16	8.1/*n*=9	35.7	34.4	19.3	5.9	4.6	729
*sop-2(bx91)**	32.3/*n*=16	11.7/*n*=25	43.1	50.1	6.8	0	0	1052
*sop-2(bx91); sor-1(RNAi)**	46.7/*n*=7	27.9/*n*=12	98.1	1.9	0	0	0	824
mes-2/mes-6(RNAi)	10.4/*n*=25	7.9/*n*=25						
sop-2(bx91); mes-2/mes-6(RNAi)†	33.2/*n*=15	12.7/*n*=30						
sop-2(bx91); mes-4(RNAi)†	29.0/*n*=16	12.1/*n*=17						

The stages of mutant larvae are determined by their size.
n, number of animals scored.
**sop-2* mutants were grown at non-permissive temperature 25°C.
†*sop-2* mutant were grown at partial non-permissive temperature 20°C. In *sor-1(RNAi)* animals, eggs laid between 0 and 4 hours showed larval lethality with some escapers developing into later larvae and adults. However, 100% of eggs laid at 4-48 hours showed larval lethality. Animals developed from the eggs laid 4-48 hours after injection were examined for Hox gene expression.

Mutations in *sor-1* and *sop-2* also cause synthetic larval lethality. Both *sor-1(RNAi)* and *sop-2* mutants are arrested at early larval stages, mostly at the L1 and L2 stages. However, 98% of the *sop-2(bx91); sor-1(RNAi)* animals are arrested at the L1 stage (Table 1). Larval arrest is not a phenotype typically associated with misregulation of Hox genes in *C. elegans*, indicating that *sor-1* and *sop-2* also function together in regulating the expression of non-homeotic genes. In these experiments, *sop-2(bx91)* appears to be a null allele and *sor-1(RNAi)* is a strong loss-of-function, if not a null, allele. The similar defects caused by mutations in *sop-2* and *sor-1* suggest that they may function redundantly in regulating some targets in common. Synergistic interactions between *sop-2* and *sor-1* suggest that they also have independent functions in Hox gene repression. They could have distinct biochemical functions within a single complex. Alternatively, they could also function in separate complexes that are involved in Hox gene regulation.

sor-1 encodes a novel protein

The *sor-1* locus was mapped by three-factor mapping between *unc-32* and *sma-3*, a small genetic interval on chromosome III (Fig. 3A). It was cloned by transformation rescue experiments (see Materials and methods). The *sor-1* rescuing DNA fragment contains a single predicted gene ZK1236.3. Most of the intron/exon boundaries were confirmed by sequencing the cDNAs of this region (Fig. 3A). We found that a 22 nucleotide leader sequence, the spliced leader 1 (SL1), is trans-spliced onto the 5′ end of *sor-1* transcripts. SL1 trans-splicing occurs in about 57% of the genes in *C. elegans* (Blumenthal and Steward, 1997). SOR-1 encodes a protein with 810 amino acids (Fig. 3B). In *sor-1(bp1)* mutants, the CAG glutamine codon at position 304 is mutated to TAG amber stop codon, resulting in the truncation of 506 amino acids in the C terminus. In *sor-1(bp2)* and *sor-1(bp3)*, an identical nonsense mutation occurs at position 634, leading to a deletion of the 176 C-terminal residues (Fig. 3B). Existence of these mutations in *sor-1* mutant alleles further confirms the identification of *sor-1*. Treatment of wild-type animals with ZK1236.3 dsRNA phenocopies *sor-1* mutants, causing the ectopic expression of *mab-5* and *egl-5* as described above. Therefore, the ectopic expression of Hox genes is due to the loss of function of *sor-1*.

SOR-1 is an RNA-binding protein

RNA has been postulated to play an important role in PcG-mediated Hox gene repression (Sun and Zhang, 2004). We therefore tested whether SOR-1 directly binds to RNA. GST-SOR-1 was purified and incubated with a radiolabeled single- or double-stranded RNA derived from the 5′UTR of the Hox gene *egl-5*, followed by electrophoretic mobility shift analysis (EMSA). The results showed that a fragment of SOR-1 binds efficiently to the RNA probes (Fig. 3C). Binding of SOR-1 to RNA is effectively competed by adding cold unrelated single- or double-stranded RNA, but not by adding tRNA or DNA (Fig. 3D). This rules out the possibility that SOR-1 is a general nucleic acid-binding protein or sticks to the charged phosphate backbone of nucleic acids non-specifically. RNA binding by SOR-1 in vitro does not appear to be sequence specific, as several unrelated RNA templates bind to SOR-1 with comparable efficiency (data not shown). However, this apparent lack of sequence specificity in vitro may not reflect the differential affinity for physiologically relevant RNA targets.

Using a series of nine overlapping SOR-1 fragments, the specific RNA-binding region of SOR-1 was mapped to an 87 amino acid region (amino acid 443 to 530). This region does not contain a recognizable RNA-binding motif, but is rich in proline and glutamine. A transgene deleted for the RNA binding region of SOR-1 cannot rescue the *sor-1* mutant phenotypes, indicating the importance of this region for *sor-1* function.

To gain further insight into the functional properties of SOR-1, we searched for putative domains in SOR-1. SOR-1 contains a region that shows weak similarity to the mouse PcG protein Rae28 (35% similarity over 381 amino acids from amino acids 115 to 496 of SOR-1) and a second region that has weak similarity to extensin 2 domain (amino acids 422 to 659 of SOR-1, $6.1e^{-01}$), which, notably, is also present in SOP-2 (amino acids 300 to 648, $5.9e^{-01}$). The extensin motif exhibits a high degree of post-translational modification, including hydroxylation, glycosylation and crosslinking, suggesting that the function of SOR-1 and SOP-2 may be regulated by post-translational modification.

SOR-1 colocalizes to the same nuclear bodies as SOP-2

To examine the expression pattern of *sor-1*, we generated transgenic lines expressing *sor-1::gfp*. The reporter gene contains the entire coding sequence and the promoter region of *sor-1*, with *gfp* inserted in the C terminus of *sor-1*. Fluorescence can be observed in all nuclei in developing embryos. At larval stages, *sor-1::gfp* becomes weaker in all of the somatic cells. Interestingly, SOR-1::GFP forms distinct nuclear bodies besides its homogenous distribution in the nucleoplasm (Fig. 4A).

To further determine the expression pattern and level of endogenous SOR-1 protein, we raised rabbit polyclonal antibodies against the N-terminal 160 amino acids of SOR-1 (see Materials and methods). The SOR-1 antibody staining pattern was consistent with those shown by the *gfp* reporter. SOR-1 is nuclear localized and

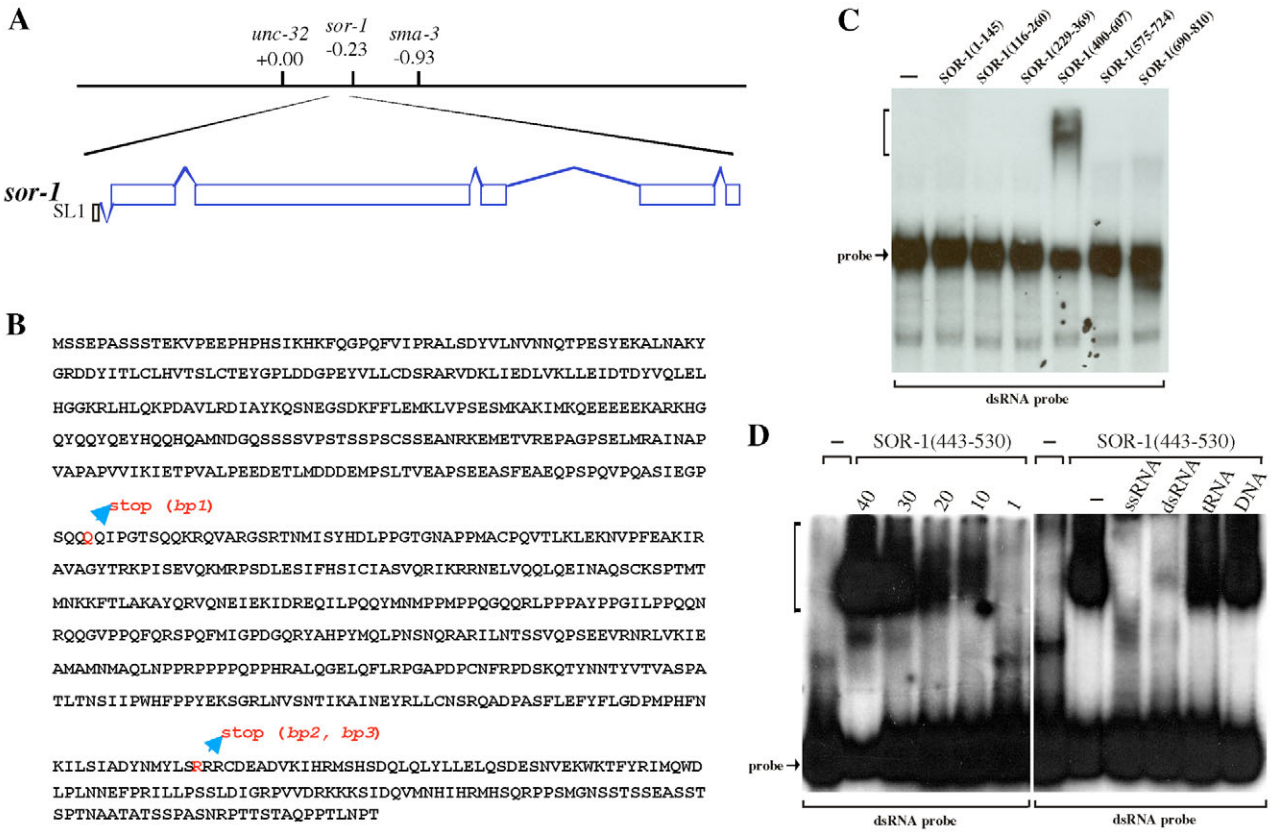

Fig. 3. Molecular structure of *sor-1* and its predicted gene product. (**A**) *sor-1* was mapped genetically to the region of LG III between *unc-32* and *sma-3*. *sor-1* is SL1 trans-spliced. The intron/exon boundaries were confirmed by sequencing the cDNAs yk526e9 and yk336g5 and the RT-PCR products. (**B**) The protein sequence of SOR-1. Mutant residues in *sor-1(bp1)*, *sor-1(bp2)* and *sor-1(bp3)* are shown in red. (**C**) Direct RNA binding by SOR-1. Various SOR-1 protein fragments were incubated with the radiolabeled RNA (5′UTR of *egl-5,* C08C3 nucleotides 41280-41840) and binding was assessed by EMSA. Protein-RNA complex is marked by bracket. (**D**) Binding of SOR-1(443-530) to RNA at a titrated protein concentration (ng/μl). Binding of RNA by SOR-1 is competed by unlabeled ssRNA (10×), dsRNA (10×) and partially by tRNA (100×), but not by DNA (10× and 100× refer to mass excess). The ssRNA, dsRNA and DNA competitors are derived from T08D10 (nucleotides 7800-8408). Similar results were obtained using several other competitors, including the ones derived from Y110A7A (nucleotides 59588-60325) and Y113G7B (nucleotides 81290-82026). Thirty ng/μl of proteins (final concentration) were used for each RNA-binding reaction unless otherwise noted for the titration experiments.

expressed in all cells from the one-cell stage onwards. As development proceeds, the expression level of SOR-1 appears to decline. From the two cell-stage onwards, SOR-1 is found to be inhomogenously expressed in the nuclei with obvious accumulations in distinct nuclear speckles (Fig. 4B). During larval development, SOR-1 is present in all somatic cell nuclei. Although at lower levels of expression, the nuclear bodies in which it localizes can still be formed.

The localization of SOR-1 in nuclear bodies lead us to investigate whether it is colocalized with SOP-2, a *C. elegans* PcG protein that is also localized in nuclear bodies, termed SOP-2 bodies (Zhang et al., 2003). Transgenic animals expressing a functional *sop-2::gfp* reporter were immunostained with anti-SOR-1 antibodies. The superimposable confocal images of SOR-1 (Fig. 4C) and SOP-2::GFP (Fig. 4D) in distinct nuclear bodies provide evidence that they are colocalized in SOP-2 nuclear bodies (Fig. 4E).

SOP-2 is required for the localization of SOR-1 into nuclear bodies

To determine the role of SOP-2 and SOR-1 in formation of the SOP-2 nuclear bodies, we examined the localization of these proteins in *sop-2* and *sor-1* mutants using anti-SOP-2 and anti-SOR-1 antibodies. Absence of SOR-1 does not affect the localization of SOP-2 (Fig. 4F,G). No obvious changes in SOP-2 localization were observed in *sor-1(RNAi)* embryos (Fig. 4F). In *sop-2(bx91)* mutant embryos, both SOP-2(bx91) (46/46 cells) and SOR-1 (45/45 cells) are localized in the cytoplasm (Fig. 4H-M), while the cytoplasmic localization of SOP-2 and SOR-1 is not evident in wild-type animals (>50 cells examined). In *sop-2(bx91)* mutant embryos, the nuclear localization of SOP-2(bx91) and SOR-1 is also seen in some cells (25/46 cells for SOP-2, and 30/45 cells for SOR-1). Thus, SOP-2 function is required for the proper localization of both SOP-2 and SOR-1.

SOR-1 and SOP-2 interact directly

Colocalization of SOR-1 and SOP-2 in SOP-2 bodies and dependence of the localization of SOR-1 on the function of SOP-2 prompted us to test whether these proteins directly interact. Glutathione-S-transferase (GST)-SOP-2 carried on glutathione sepharose beads was incubated with S^{35} labeled SOR-1. The retained proteins after several washes were examined by gel electrophoresis. We found that SOP-2 specifically binds to the N-terminal 503 amino acids of SOR-1 (Fig. 5A). And, vice versa, GST-SOR-1 interacts with the ^{35}S-labeled SOP-2 (Fig. 5B). As SOP-2 contains an RNA-binding activity and reactions of in vitro translated

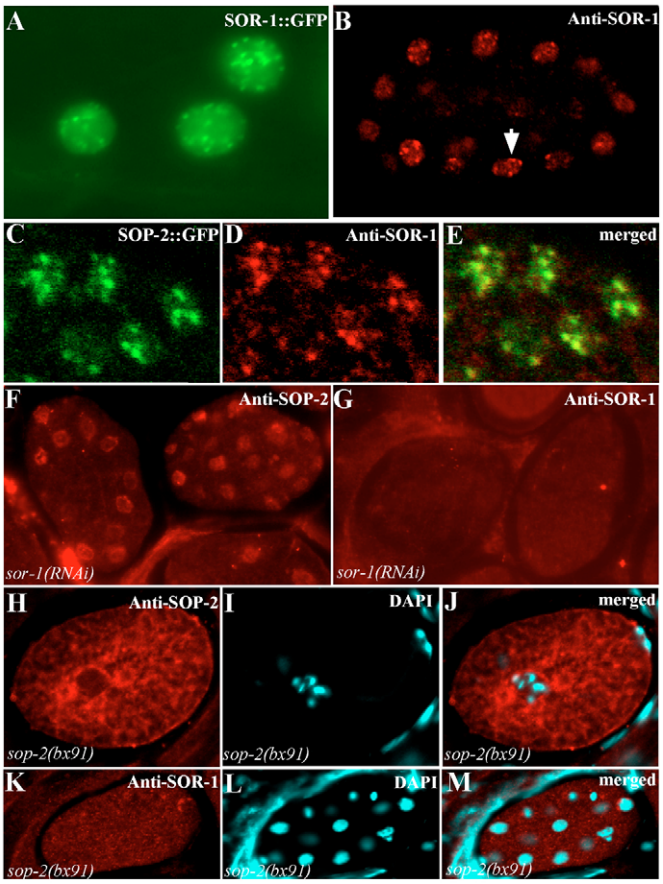

Fig. 4. Colocalization of SOR-1 and SOP-2 in nuclear bodies. (**A**) Expression of *sor-1::gfp* in the hypodermal nuclei of adult animals. SOR-1::GFP is distributed throughout the nuclei with accumulation in distinct nuclear speckles. (**B**) Expression of SOR-1 in an early embryo. Polyclonal antibodies against SOR-1 coupled with rhodamine-conjugated goat anti-rabbit IgG were used to detect expression. SOR-1 is localized in distinct nuclear bodies (arrow). (**C-E**) Localization of SOP-2::GFP (C) and SOR-1 (D) into nuclear bodies. (E) The superimposable confocal images of SOR-1 (red) and SOP-2::GFP (green) shows that they colocalize (yellow). (**F**,**G**) Unaltered localization of SOP-2 into nuclear bodies in *sor-1(RNAi)* mutant embryos (F). Absence of staining of SOR-1 in *sor-1(RNAi)* mutant embryos confirms the specificity of the antibody (G). (**H-J**) The merged images of SOP-2 (H) and DAPI (I) shows that SOP-2(bx91) is localized in cytoplasm in *sop-2(bx91)* mutant embryos (J). (**K-M**) The merged images of SOR-1 (K) and DAPI (L) shows that SOR-1 is localized in cytoplasm in *sop-2(bx91)* mutant embryos (M). (B-M) Single confocal sections.

proteins contain many RNAs, we further tested whether the interaction between SOR-1 and SOP-2 was RNA dependent. We found the interaction was not significantly reduced after the reaction was treated with RNase (data not shown), supporting the notion that SOP-2 and SOR-1 interact directly.

To map the minimal interacting domain in each protein, a series of GST fusion fragments of SOR-1 and SOP-2 were used in GST pull-down assay in vitro. The SOR-1 interacting domain in SOP-2 is mapped to its N terminus, rather than to its C-terminal protein-protein interacting SAM domain, which mediates the self-association of SOP-2 (Fig. 5A,C). However, the SOP-2 interacting domain in SOR-1 is located to its N terminus, which is predicted to be rich in helical structures (Fig. 5B,C).

Interaction between the N terminus of SOP-2 and SOR-1 was further demonstrated by the co-purification of SOP-2 and SOR-1 expressed in bacterial. GST-SOP-2 (amino acids 58-140) and non-tagged SOR-1 (amino acids 48-200) were co-expressed in *E. coli* strain BL21. SOP-2 proteins were purified by the GST affinity column followed by ion exchange and Superdex columns. The elute was subjected to SDS-PAGE and visualized by Coomassie staining. We found that SOR-1 was specifically co-purified with SOP-2 (Fig. 5D).

SOP-2 synergistically interacts with the MES-2/MES-6 complex in regulating non-Hox genes

In addition to a role in germline development (Fong et al., 2002), the MES-2/MES-6 complex is also involved in transcriptional repression of Hox genes (Ross and Zarkower, 2003). To determine whether the MES-2/MES-6 complex and the putative SOP-2/SOR-1 complex function together in repression of Hox genes, we compared the expression pattern of Hox reporter genes in *sop-2(bx91); mes-2/mes-6(RNAi)* mutants with those in single mutants. As previously reported, *egl-5::gfp* is weakly ectopically expressed in *mes-2/mes-6(RNAi)* mutants (Table 1). However, no synergistic interactions are observed between *sop-2(bx91)* and *mes-2/mes-6 (RNAi)* (Table 1). The onset of ectopic expression and the expression domains of *egl-5::gfp* in *sop-2(bx91); mes-2/mes-6(RNAi)* is similar to those in *sop-2* single mutants. Lack of synergy between mutations in these two complexes and the weaker effects of *mes-2* and *mes-6* on Hox genes suggest that the MES-2/MES-6 complex is unlikely to be a prerequisite to recruit the SOP-2/SOR-1 complex to Hox genes.

However, *sop-2(bx91)* and a *mes-6* mutation cause synthetic lethality and a bursting vulva phenotype. Shifting *sop-2; mes-6* L1/L2 larvae from 15°C to the semi-permissive temperature 20°C, 69.3% of animals (n=322) died of bursting vulva, whereas none of the single mutants (n>100) has a bursting vulva phenotype. Shifting *sop-2; mes-6* mutant eggs laid in the first 8 hours from 15°C to non-permissive temperature 23°C, 100% animals arrested at L1 stage, while *sop-2(bx91)* mutants alone arrested at L2 to L4 stages and all *mes-6* mutant animals developed into adults. No such synergistic interactions have been observed between *sop-2* and a *mes-4* mutation, consistent with the fact that MES-4 is not a component of the MES-2/MES-6 complex. These synergistic interactions suggest that SOP-2/SOR-1 and MES-2/MES-6 could act in a single pathway or their effects in separate pathways may be cumulative in regulating non-Hox genes.

SOR-1 and SOP-2 orthologs are absent from other organisms

SOR-1 and SOP-2 do not appear to be homologs of known PcG proteins in other organisms, although several properties of the SOR-1/SOP-2 complex, including localization into nuclear bodies and binding to RNA, are reminiscent of the ones of the PRC1 complex in other organisms. To determine whether SOR-1/SOP-2 defines a new global Hox gene repression system, we examined whether other organisms contain homologs of SOR-1 and SOP-2. Surprisingly, no recognizable SOR-1 or SOP-2 homologs were identified in fly, mouse and human by searching available databases.

To determine whether SOR-1 and SOP-2 are conserved among nematodes, we searched the sequence of *C. briggsae*, which is the closest known species to *C. elegans* and has almost identical morphology. The *C. briggsae* sequence is currently more than 98% complete and a genome-wide comparison of *C. elegans* and *C. briggsae* reveals that 89% of the *C. briggsae* genes have either

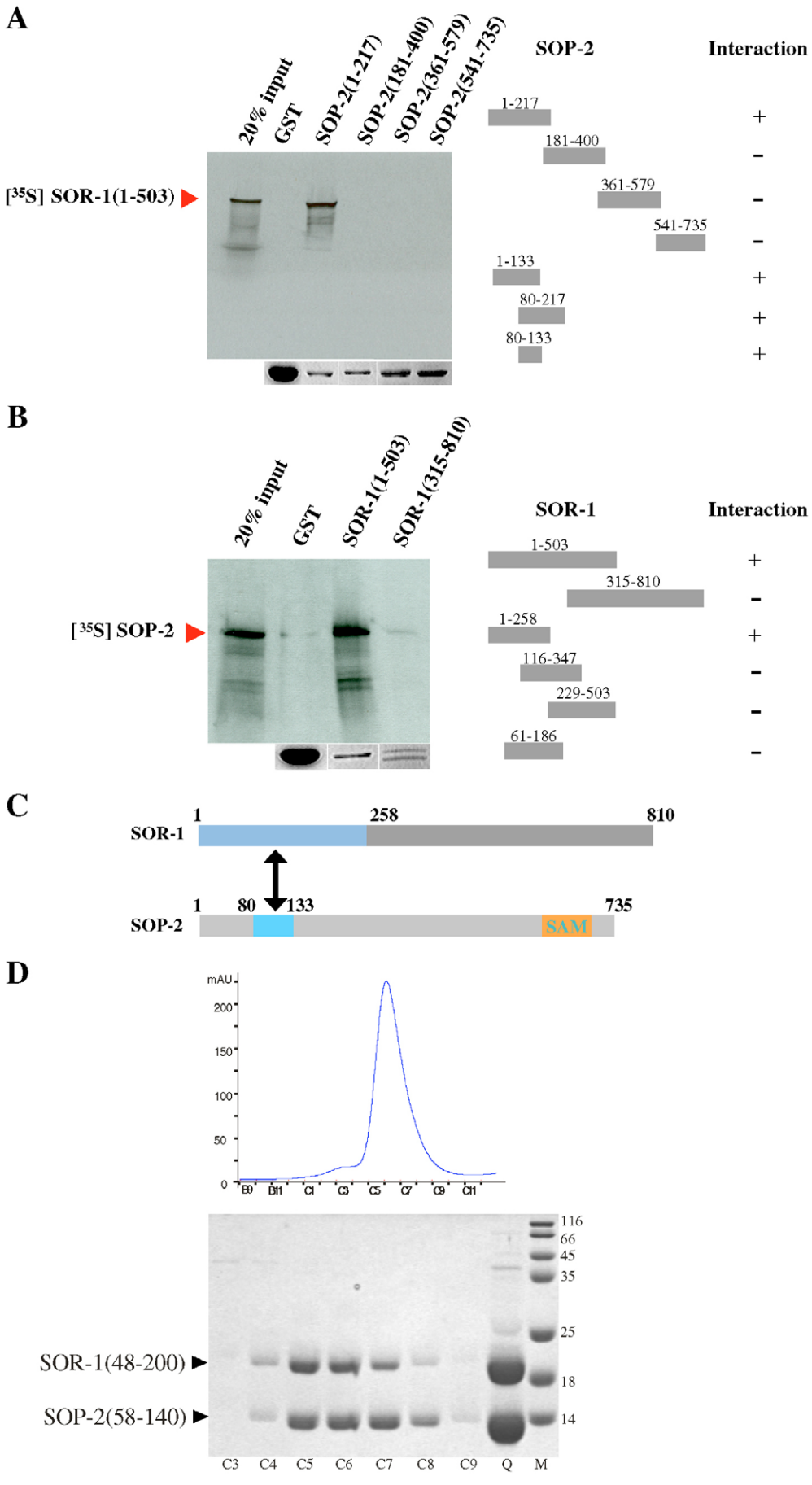

Fig. 5. Direct interaction between SOR-1 and SOP-2. (**A**) GST-SOP-2 fusion proteins were incubated with ^{35}S-labeled SOR-1 (amino acids 1-503) to map the interaction domain in SOP-2. Only the N-terminal domain (amino acids 1-217) of SOP-2 bound SOR-1. Smaller fragments of SOP-2 were used to further map the minimal interaction domain. (**B**) The N terminus of SOR-1 interacts with SOP-2. GST-SOR-1 fusion proteins were incubated with ^{35}S-labeled SOP-2. Twenty percent of the ^{35}S-labeled protein used in the binding reactions was shown. No interaction was detected between the ^{35}S-labeled protein and GST alone or GST fused to the C terminus of SOR-1 but binding was detected between SOP-2 and the N terminus of SOR-1. Smaller fragments of SOR-1 were used to map the interaction domain further. (**C**) Schematic representation of protein interaction domains of SOR-1 and SOP-2 (connected by arrow). SOP-2 also contains a protein interaction SAM domain, which mediates its self-association (Zhang et al., 2003). (**D**) Interaction between N terminus of SOP-2 (residues 58-140) and SOR-1 (residues 48-200). The chromatogram of gel filtration is shown on the top and the fractions from the complex are shown at the bottom. Q indicates proteins purified through ion exchange column.

orthologs (62%) or multiple matches (27%) in the *C. elegans* genome (Stein et al., 2003). As in *C. elegans*, the ESC/E(Z) complex is present in *C. briggsae*. However, *C. briggsae* lacks PRC1 complex genes. Surprisingly, orthologs of SOR-1 or SOP-2 could not be identified in *C. briggsae* genome. The most closely related protein to SOR-1 in *C. briggsae* is CBP23695 (encoded by CBG18142) ($2.6e^{-10}$, 45.3%) and to SOP-2 is CBP06294 (encoded by CBG02561) ($1.1e^{-19}$, 52.4%). However, these two proteins do not appear to be the orthologs of SOR-1 and SOP-2. Regions of homology are restricted to the low complexity parts of SOR-1 or SOP-2. More importantly, functionally significant domains in SOR-1 (e.g. the RNA binding domain) and SOP-2 (e.g. the SAM domain and the RNA-binding domains) are not present in these proteins. In addition, CBP06294 has a clear ortholog in *C. elegans*, the gene ZK84.1 ($7.4e^{-175}$, 82.1%), and CBP23695 is closer to the *C. elegans* gene H20J18.1a ($3.7e^{-16}$, 54.6%). Interestingly, orthologs of the genes adjacent to both *sor-1* and *sop-2* do exist in *C. briggsae* and the organization of these neighboring genes in *C. briggsae* is comparably co-linear with the ones in *C. elegans* (Fig. 6A,B). For example, as in *C. elegans*, there are two genes between CBG18138 and CBG18143, and one gene between CBG20942 and CBG20938 in *C. briggsae*. This further supports the assertion that SOR-1 and SOP-2 orthologs are absent from *C. briggsae*. In conclusion, the putative SOR-1/SOP-2 complex, the function of which is similar to that of the PcG complex in other organisms in maintaining the repressed state of Hox genes, may have undergone specific evolution in *C. elegans*, even though the underlying mechanisms appear to be conserved.

DISCUSSION

The role of SOP-2 and SOR-1 in Hox gene repression

Despite a lineage-driven mode of development in *C. elegans*, we provide evidence that *sor-1* and *sop-2* function as PcG-like genes in global transcriptional repression of Hox gene. As in *PcG* mutants in other organisms, the expression of Hox genes is correctly initiated in *sor-1* and *sop-2* mutants. However, the repressed state of Hox genes fails to be maintained outside their normal expression domains during subsequent development. SOR-1 and SOP-2 may form a complex in Hox gene repression. Both GFP reporters and antibody staining indicate that SOR-1 and SOP-2 are colocalized in nuclear bodies and the localization of SOR-1 depends on SOP-2. Furthermore, SOR-1 and SOP-2 directly interact in vitro. The synergistic phenotypes of *sop-2; sor-1* double mutants further support the conclusion that SOP-2 and SOR-1 function together in gene regulation. In other organisms, combination of mutations in different *PcG* genes typically produces more than additive effects (Jurgens, 1985). These synergistic effects could be due to the disruption of the function or integrity of the complex by a cumulative effect of two simultaneously mutations.

Although *sor-1* and *sop-2* work in concert in Hox gene repression, they do not appear to make an equal contribution to gene regulation. Mutations in *sop-2* cause more dramatic effects than those in *sor-1* mutants, demonstrated by a greater extent and degree of ectopic expression of an *egl-5* reporter. Furthermore, the expression patterns of *sor-1* and *sop-2* are not identical. The expression level of *sor-1* is weaker than that of *sop-2*, especially in the post-embryonic somatic cells. At later embryonic and larval stages, the majority, if not all, of SOP-2 is localized in nuclear bodies. Although SOR-1 is colocalized with SOP-2 in bodies, it also diffuses in the nuclei at all stages. Thus, SOR-1 is likely to be needed only in specific cell types and developmental stages. Differences in the timing and tissue specificity of homeotic misexpression have also been observed in *PcG* mutants in other organisms (Soto et al., 1995). Homeotic defects are more severe in the loss of function of *Pc*, *Ph* and *Scm* than those in loss of function of other *PcG* components. This could be due to their distinct roles in establishing repressive chromatin structures. For example, PSC is essential for compacting the nucleosomal arrays, while RING1 mediates the ubiquitination of H2A (Francis et al., 2004; de Napoles et al., 2004; Wang et al., 2004).

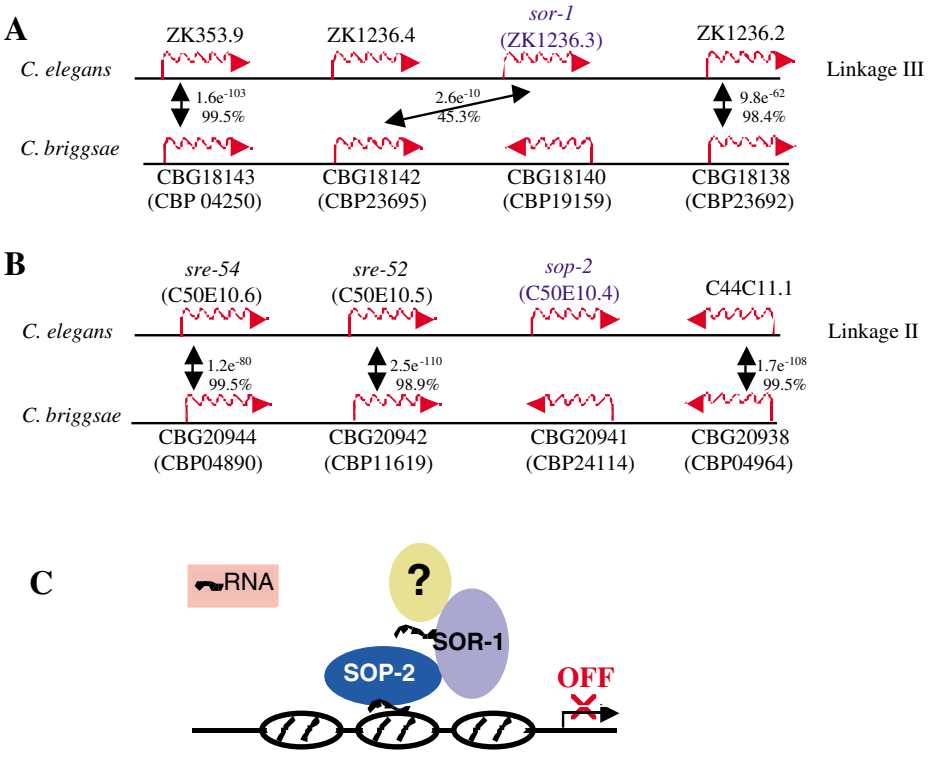

Fig. 6. Lack of orthologs of SOR-1 and SOP-2 in *C. briggsae*. (**A**) Schematic representation of *sor-1* and its neighboring genes in *C. elegans* and in *C. briggsae*. ZK1236.4 corresponds to a transposon in the genome. CBG18140 does not have a clear homolog in *C. elegans*. Although *sor-1* (ZK1236.3) shows some similarity with CBG18142, the latter does not appear to be the ortholog of *sor-1* (see Results). (**B**) Schematic representation of *sop-2* and its neighboring genes in *C. elegans* and in *C. briggsae*. CBG20941 does not have a clear homolog in *C. elegans*. *sre-52* and *sre-54* encode a class of 7 TM chemoreceptors (*sre* family). (**C**) Model for the putative SOP-2/SOR-1 complex-mediated transcriptional repression. SOR-1, SOP-2 and perhaps other unidentified components form an RNA-binding complex, which may further recruit chromatin remodeling activity. The RNA components could be involved in targeting the SOP-2/SOR-1 complex to target loci and/or maintaining the integrity of the complex and the distinct nuclear bodies formed by SOR-1 and SOP-2.

Other possible integral components of the SOP-2/SOR-1 complex are RNAs. Both SOR-1 and SOP-2 contain RNA-binding activity. Having two RNA-binding proteins involved in the same process supports the role of RNA in epigenetic silencing of Hox genes, although the nature of the RNA components in the SOP-2/SOR-1 complex has yet to be identified. siRNAs and structural *roX1* and *roX2* RNAs are important for the formation and targeting of the RITS complex and the *Drosophila* dose compensation complex (DCC), respectively (Verdel et al., 2004; Motamedi et al., 2004; Akhtar, 2003). Analogous to other RNA-mediated epigenetic gene regulatory phenomena, the RNA components could be involved in maintaining the integrity of the PcG complex and/or targeting the PcG complex to specific loci (Fig. 6C).

Transcriptional repression mediated by the SOP-2/SOR-1 complex and the PRC1 complex

In other organisms, the ESC/E(Z) complex and the PRC1 complex function together in the maintenance of the Hox gene expression. The ESC/E(Z) complex may be involved in recruiting the PRC1 complex to the target loci. Unlike in *sop-2* and *sor-1* mutants, which have widespread ectopic expression of Hox genes, the effect of *mes-2* and *mes-6* mutations on Hox gene repression is subtle and is only apparent in sensitized genetic backgrounds (Ross and Zarkower, 2003). Lack of synergy between *sop-2(bx91)* and *mes-2/mes-6(RNAi)* further supports the conclusion that the MES-2/MES-6 complex may not be involved in recruiting the SOP-2/SOR-1 complex in Hox gene repression.

In spite of a lack of obvious sequence similarity, several conserved properties of the PRC1 and putative SOP-2/SOR-1 complexes suggest that they appear to employ a conserved mechanism in Hox gene repression. First, components of the PRC1 complex and the SOP-2/SOR-1 complex are localized into distinct nuclear bodies, although the mechanistic role of the bodies in Hox gene repression remains unknown. Second, the protein-protein interaction SAM domain is present in both complexes. Domain-swapping experiments indicate that only the SAM domains of PcG proteins, including PH and SCM, but not the ones of non-PcG proteins, including TEL and L(3)MBT, can functionally substitute for the SAM domain of SOP-2 in terms of repression of Hox genes and proper formation of SOP-2 nuclear bodies (Zhang et al., 2004a). Overexpression of the SAM domain behaves as a dominant-negative regulator that compromises the function of the SOP-2/SOR-1 complex and the PRC1 complex in Hox gene repression (Y.S. and H.Z., unpublished) (Peterson et al., 2004). These observations indicate an important role of the SAM domain in PcG-mediated Hox gene repression. Third, both of the SOP-2/SOR-1 complex and the PRC1 complex contain RNA-binding activity (Zhang et al., 2004b). Although a role for RNA in PcG-mediated gene silencing of Hox genes in other organisms has not been demonstrated, PcG proteins are involved in some RNA-dependent silencing processes, such as X chromosome inactivation and silencing of transgenes in fly (Plath et al., 2003; Plath et al., 2004; Pal-Bhadra et al., 2002). The evolutionary constraint on the PRC1 complex is, thus, likely to be conferred at the mechanistic level but not at the protein level. The reason for the maintenance of the PRC1 complex over very large evolutionary distance encompassing insects and vertebrates, while in nematodes a similar complex appears to be evolutionarily volatile, remains to be explained.

We thank Drs Xiaodong Wang, Scott Emmons, Robyn Lints, Rachel Palmer and members in our laboratory for their helpful comments on the manuscript. We also thank Drs Gang Zhi and Zhifu Han for the fusion proteins and the generation of SOR-1 antibodies. This work was supported by the National High Technology Projects 863.

References

Akhtar, A. (2003). Dosage compensation: an intertwined world of RNA and chromatin remodelling. *Curr. Opin. Genet. Dev.* **13**, 161-169.

Albertson, D. G. (1984). Localization of the ribosomal genes in *Caenorhabditis elegans* chromosomes by in situ hybridization using biotin-labeled probes. *EMBO J.* **3**, 1227-1234.

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

Blumenthal, T. and Steward, K. (1997). RNA processing and gene structure. In *C. elegans II* (ed. D. Riddle et al.), pp. 117-147. New York: Cold Spring Harbor Laboratory Press.

de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M. et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663-676.

Emmons, S. W. (1999). Cell fate determination in *Caenorhabditis elegans* ray develpment. In *Cell Lineage and Fate Determination* (ed. S. A. Moody), pp. 139-155. San Diego: Academic Press.

Ferreira, H. B., Zhang, Y. H., Zhao, C. and Emmons, S. W. (1999). Patterning of *Caenorhabditis elegans* posterior structures by the *Abdominal-B* homolog, *egl-5*. *Dev. Biol.* **207**, 215-228.

Fong, Y., Bender, L., Wang, W. and Strome, S. (2002). Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of *C. elegans*. *Science* **296**, 2235-2238.

Francis, N. J., Kingston, R. E. and Woodcock, C. L. (2004). Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574-1577.

Gellon, G. and McGinnis, W. (1998). Shaping animal body plans in development and evolution by modulation of Hox expression patterns. *BioEssays* **20**, 116-125.

Heard, E. (2004). Recent advances in X-chromosome inactivation. *Curr. Opin. Cell Biol.* **16**, 247-255.

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

Holdeman, R., Nehrt, S. and Strome, S. (1998). MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a *Drosophila* Polycomb group protein. *Development* **125**, 2457-2467.

Jurgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* **316**, 153-155.

Kenyon, C. J., Austin, J., Costa, M., Cowing, D. W., Harris, J. M., Honigberg, L., Hunter, C. P., Maloof, J. N., Muller-Immergluck, M. M., Salser, S. J. et al. (1997). The dance of the Hox genes: Patterning the anteroposterior body axis of *Caenorhabditis elegans*. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 293-305.

Korf, I., Fan, Y. A. and Strome, S. (1998). The Polycomb group in *Caenorhabditis elegans* and maternal control of germline development. *Development* **125**, 2469-2478.

Levine, S. S., King, I. F. and Kingston, R. E. (2004). Division of labor in polycomb group repression. *Trends Biochem. Sci.* **29**, 478-485.

Lippman, Z. and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. *Nature* **431**, 364-370.

Lund, A. H. and van Lohuizen, M. (2004). Polycomb complexes and silencing mechanisms. *Curr. Opin. Cell Biol.* **16**, 239-246.

Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P. and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**, 789-802.

Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E. and Simon, J. A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* **111**, 197-208.

Netter, S., Faucheux, M. and Theodore, L. (2001). Developmental dynamics of a polyhomeotic-EGFP fusion in vivo. *DNA Cell Biol.* **20**, 483-492.

Pal-Bhadra, M., Bhadra, U. and Birchler, J. A. (2002). RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in *Drosophila*. *Mol. Cell* **9**, 315-327.

Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Rao, M., Bhadra, U., Birchler, J. A. and Elgin, S. C. (2004). Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**, 669-672.

Panning, B. and Jaenisch, R. (1998). RNA and the epigenetic regulation of X chromosome inactivation. *Cell* **93**, 305-308.

Peterson, A. J., Mallin, D. R., Francis, N. J., Ketel, C. S., Stamm, J., Voeller, R. K., Kingston, R. E. and Simon, J. A. (2004). Requirement for sex comb on midleg protein interactions in *Drosophila* polycomb group repression. *Genetics* **167**, 1225-1239.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B. and Zhang, Y. (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300**, 131-135.

Plath, K., Talbot, D., Hamer, K. M., Otte, A. P., Yang, T. P., Jaenisch, R. and

Panning, B. (2004). Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J. Cell Biol.* **167**, 1025-1035.
Ross, J. M. and Zarkower, D. (2003). Polycomb group regulation of Hox gene expression in *C. elegans*. *Dev. Cell* **4**, 891-901.
Salser, S. J. and Kenyon, C. (1996). A *C. elegans* Hox gene switches on, off, on and off again to regulate proliferation, differentiation and morphogenesis. *Development* **122**, 1651-1661.
Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P. and Brockdorff, N. (2003). Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev. Cell* **4**, 481-495.
Soto, M. C., Chou, T. B. and Bender, W. (1995). Comparison of germline mosaics of genes in the Polycomb group of *Drosophila melanogaster*. *Genetics* **140**, 231-243.
Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A. et al. (2003). The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* **1**, E45.
Sun, Y. Y. and Zhang, H. (2004). A unified mode of epigenetic gene silencing. *RNA Biol.* **1**, 132-134.
van der Vlag, J. and Otte, A. P. (1999). Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nat. Genet.* **23**, 474-478.
Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I. and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672-676.
Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833-1837.
Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S. and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878.
Zhang, H., Azevedo, R. B., Lints, R., Doyle, C., Teng, Y., Haber, D. and Emmons, S. W. (2003). Global regulation of Hox gene expression in *C. elegans* by a SAM domain protein. *Dev. Cell* **4**, 903-915.
Zhang, H., Smolen, G. A., Palmer, R., Christoforou, A., Van den Heuvel, S. and Haber, D. A. (2004a). SUMO modification is required for *in vivo* Hox gene regulation by the *C. elegans* Polycomb group protein SOP-2. *Nat. Genet.* **36**, 507-511.
Zhang, H., Christoforou, A., Aravind, L., Emmons, S. W., van den Heuvel, S. and Haber, D. A. (2004b). The *C. elegans Polycomb* gene *sop-2* encodes a RNA binding protein. *Mol. Cell* **14**, 841-847.