Somatic sex determination in *Caenorhabditis elegans* is modulated by SUP-26 repression of *tra-2* translation

James Mapes*, Jeng-Ting Chen†, Jaou-Song Yu†, and Ding Xue*†

*Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309; and †Graduate Institute of Biomedical Sciences, College of Medicine, Chang Gung University, Tao-Yuan 333, Taiwan

Edited by Judith Kimble, University of Wisconsin, Madison, WI, and approved September 10, 2010 (received for review April 12, 2010)

Translational repression mediated by RNA-binding proteins or micro RNAs has emerged as a major regulatory mechanism for fine-tuning important biological processes. In *Caenorhabditis elegans*, translational repression of the key sex-determination gene *tra-2* (tra, transformer) is controlled by a 28-nucleotide repeat element, the TRA-2/GLI element (TGE), located in its 3′ untranslated region (UTR). Mutations that disrupt TGE or the germline-specific TGE-binding factor GLD-1 increase TRA-2 protein expression and inhibit sperm production in hermaphrodites. Here we report the characterization of the *sup-26* gene, which regulates sex determination in the soma and encodes an RNA recognition motif (RRM)-containing protein. We show that *SUP-26* regulates the level of the TRA-2 protein through TGE in vivo and binds directly to TGE in vitro through its RRM domain. Interestingly, *SUP-26* associates with poly(A)-binding protein 1 (PAB-1) in vivo and may repress *tra-2* expression by inhibiting the translation-stimulating activity of PAB-1. Taken together, our results provide further insight into how mRNA-binding factors repress translation and modulate sexual development in different tissues of *C. elegans*.

Translational repression through cis-acting elements in mRNAs is an important post-transcriptional regulatory mechanism in numerous biological systems (1). Analysis of the lengths of 5′ and 3′ untranslated regions (UTRs) shows that the average lengths of 5′ UTRs are relatively constant across phyla, whereas the lengths of 3′ UTRs increase with organism complexity (200 bp for yeast and 500 bp for humans), suggesting that they may be more highly regulated during animal development (2). Sequence elements or modifications in 3′ UTR are known to control the subcellular localization, stability, and translational efficiency of mRNAs. For example, the poly(A) sequence is important not only for the stability of mRNAs, but also for stimulating translation initiation by facilitating interaction of poly(A)-binding (PAB) protein with translation initiation factor eIF4G at the 5′ cap and formation of circularized mRNA (1). Moreover, numerous important developmental regulators, such as Bicoid and the cytoplasmic polyadenylation-element binding protein, act by binding 3′ UTRs and repressing translation (1). Finally, translational repression by microRNAs (miRNAs) is mediated primarily by formation of nonperfect duplexes between miRNAs and their mRNA targets at 3′ UTRs, which induces the formation of the translation repressive complex termed the RNA induced silencing complex (3). Therefore, cis elements in the 3′ UTR of mRNAs play critical roles in regulating the efficiency of translation.

Sex differentiation in *Caenorhabditis elegans* is determined by the X chromosome:autosome ratio: 1:2 results in XO males and 1:1 results in XX hermaphrodites (4, 5). Hermaphrodites are essentially females that produce sperm before oogenesis and are capable of self-fertilization and mating with other males. Male development is initiated by expression of a male-promoting secreted protein, HER-1 (hermaphrodization) (6, 7), which binds and inactivates the hermaphrodite-promoting transmembrane receptor TRA-2 (TRA, transformer), which is also important for sperm production (8, 9). TRA-2 interacts with and suppresses the male-promoting activity of an intracellular protein complex containing FEM-1 (feminization), FEM-2, and FEM-3 (10–13). How TRA-2 inhibits the activities of FEM proteins is poorly understood, but it may involve cleavage of the intracellular domain of TRA-2 by the TRA-3 calpain protease and subsequent translocation of the TRA-2 intracellular domain to the nucleus (14–16). The FEM-1/FEM-2/FEM-3 complex promotes male development by inhibiting the activity of the terminal sex-determination factor, TRA-1A, a zinc-finger transcription factor that promotes hermaphrodite development by repressing expression of genes required for sperm production and somatic male development (17, 18).

Translation repression plays an important role in regulating *C. elegans* sex differentiation. For example, the activities of *fem-3* and *tra-2* are regulated by translational repressors acting in specific tissues (19, 20). Several gain-of-function, feminizing mutations in the *tra-2* gene were found to alter the *tra-2* 3′ UTR (20, 21). Further studies revealed that two different sequence elements in the *tra-2* 3′ UTR regulate TRA-2 translation. First, the TRA-2 retention element retains the *tra-2* message in the nucleus and thus prevents translation (22). Second, the TRA-2/GLI element (TGE) is a conserved 28-nucleotide repeat element found in both *C. elegans* *tra-2* and *D. melanogaster* GLI3′ UTRs (20, 23). Mutations disrupting TGEs increase *tra-2* poly(A) tail length (23, 24) and TRA-2 protein levels in both the germline and the soma (23, 25), suggesting that TGEs negatively regulate *tra-2* expression. In the germline, repression of *tra-2* translation is mediated by GLD-1 (germline development defective), a TGE-binding protein and a member of the STAR family of RNA-binding proteins (25), and by FOG-2 (feminization of germline), a GLD-1–interacting and F-box–containing protein (25, 26). FOG-2, GLD-1, and *tra-2* 3′ UTR form a ternary complex to repress *tra-2* translation in the germline (26). However, GLD-1 and FOG-2 are expressed only in the germline, and it is unclear how TGEs mediate repression of *tra-2* translation in somatic tissues (26, 27).

Here we report the molecular and biochemical characterization of *sup-26* (suppressor). Loss-of-function (lf) mutations in *sup-26* are semidominant suppressors of the masculinization defect in her-1(n693gf) XX animals and can suppress other masculinization defects in the absence of *her-1*, indicating that *sup-26* likely acts downstream of *her-1* to affect somatic sex determination (28). We find that *sup-26* encodes an RNA recognition motif (RRM) containing protein that is expressed widely in somatic tissues, regulates the level of the *tra-2* protein in the soma through the TGEs in the *tra-2* 3′ UTR, and binds directly to TGEs in vitro. Therefore, *SUP-26* is a somatic TGE-
binding factor that promotes male development by repressing \textit{tra-2} translation.

**Results and Discussion**

**sup-26 Encodes a Protein with Two RRM Motifs.** \textit{sup-26} (\textit{n1091}) was isolated as a semidominant suppressor of the masculinized defect of \textit{her-1(n695g)} XX animals (Table S1) and mapped to a small genetic interval at approximately -3.2 genetic map units on linkage group (LG) III (28). To clone \textit{sup-26}, fosmids covering this genetic interval were injected into \textit{sup-26(n1091)}; \textit{her-1(n695g)} animals and tested for restoration of the masculinized \\textit{Tra} phenotype (Fig. 1A). Two overlapping fosmids, \textit{W DRM0627E08} and \textit{WRM066ddA01}, each restored the \textit{Tra} phenotype (Fig. 1A). A translational GFP fusion that contains a 4-kb genomic fragment, including a 1.1-kb sequence 5' of the \textit{R10E4.2} start codon, also restored the \textit{Tra} phenotype in \textit{sup-26} (\textit{n1091}); \textit{her-1(n695g)} animals (Table S1; Methods), indicating that \textit{R10E4.2} is responsible for the rescuing activity. We determined \textit{R10E4.2} DNA sequences from two different \textit{sup-26} mutants, \textit{sup-26(ct49)} and \textit{sup-26(n1091)}, and found a C-to-T transition in \textit{R10E4.2} (Fig. 1A), which converts codon Q20 to an ochre stop codon, and a G-to-A transition in \textit{sup-26(n1091)}, which converts codon C215 to a tyrosine codon. Two independently isolated deletion mutations, \textit{gk403} (a 424-bp deletion) and \textit{gk426} (a 676-bp deletion), each of which removes the first two exons of \textit{R10E4.2} (Fig. 1A), also suppressed the \textit{her-1(n695g)} Tra phenotype, confirming that \textit{R10E4.2} is \textit{sup-26} (Table S1). Given the molecular nature of the two deletions and \textit{sup-26(ct49)}, they are likely strong \textit{if} or null mutations. However, \textit{sup-26} mutant males or hermaphrodites alone display no obvious defect in sex determination (Tables S1 and S2). Therefore, \textit{sup-26} appears to be a modulator of the sex determination pathway, fine-tuning the pathway to ensure appropriate sexual development.

We performed reverse transcription PCR amplification (RT-PCR) with primers corresponding to the predicted 5' and 3' ends of the \textit{sup-26} coding sequence (http://www.wormbase.org/) and identified two distinct transcripts, \textit{sup-26a} and \textit{sup-26b}, which encode 357 and 409 amino acid products, respectively (Fig. 1B). The predicted products of both transcripts contain two RRMs that share 77% and 74% sequence similarity to the consensus RRM sequence, respectively (Fig. 1B), suggesting that \textit{SUP-26} may bind RNA. When expressed under the control of the \textit{sup-26} promoter, each of the transcripts masculinized \textit{sup-26(n1091)}; \textit{her-1(n695g)} animals (Fig. 1B), indicating that both \textit{sup-26} isoforms are functional.

**sup-26 is Broader Expression in Somatic Cells and Localizes to the Cytoplasm.** To determine where \textit{SUP-26} might function, we examined the expression pattern of the \textit{SUP-26::GFP} translational fusion, which fully rescued the \textit{sup-26(n1091)} phenotype (Table S1). We found that \textit{SUP-26::GFP} was expressed in most, if not all, somatic cells, starting from the early gastrula through adulthood. \textit{SUP-26::GFP} localized to the cytoplasm and was largely excluded from the nucleus (Fig. 2). There was no apparent difference in \textit{SUP-26::GFP} expression patterns between male and hermaphrodite \textit{L4} larvae or adults (Fig. 2B). Based on data from the Nematode Expression Pattern Database (http://nematode.lab.nig.ac.jp), in situ hybridization experiments using either \textit{sup-26a} or \textit{sup-26b} cDNA as probes reveal that the \textit{sup-26} messages are absent from early meiotic-stage germ cells but are present in oocytes.

**TRA-2 Protein Expression Is Increased by \textit{sup-26} Loss-of-Function Mutations.** Previous genetic analysis indicates that \textit{sup-26} may regulate sexual development through \textit{tra-2} (28). We thus examined whether \textit{sup-26} mutations affect \textit{tra-2} gene expression. We generated a 4-kb transgene that contains the entire \textit{tra-2} operon (\textit{ppp-1} and \textit{tra-2}), including an 816-bp promoter upstream of \textit{ppp-1}, the first gene of the operon, the coding region of \textit{ppp-1}, the \textit{tra-2} coding region fused at its carboxyl terminus with GFP or \textit{fut} RNA. When expressed under the control of the \textit{tra-2} promoter, \textit{TRA-2::GFP} was observed exclusively in the nucleus as previously described (Fig. 3B) (15). Interestingly, \textit{TRA-2::GFP} was expressed at higher levels in \textit{sup-26(gk426)} animals than in wild-type animals (Fig. 3B).

**A.** Cloning of \textit{sup-26}. (A) Fosmids used in \textit{sup-26(n1091)} rescue experiments and their relative base-pair positions on LGII are shown. Transgenic \textit{sup-26(n1091)}; \textit{her-1(n695g)} animals carrying the indicated fosmid DNA as extrachromosomal arrays were generated and scored for restoration of the masculinized (\textit{Tra}) phenotype as described in Methods. The number of rescued lines vs. total lines generated are indicated at the right. ORFs in the overlapping region of two rescuing fosmids (WRM066ddA01 and WRM066ddE03) are indicated, with boxes representing exons and lines representing intronic sequences. The positions of the \textit{ct40} and \textit{n1091} mutations are indicated by arrow. Two deletion alleles (\textit{gk403} and \textit{gk426}) and the \textit{sup-26} regions removed by these mutations are represented below the \textit{sup-26} ORF. (B) A schematic of \textit{sup-26} transcripts and alignment of the consensus RRM (accession no. PF00076) with the two \textit{SUP-26} RRMs. Expression of these two transcripts under the control of the \textit{sup-26} promoter rescued the \textit{sup-26(n1091)} phenotype. Uppercase letters indicate the most conserved residues of \textit{RRMs}. The middle rows show residues that are identical (\textit{lett}) or conservative changes (\textit{con}). The \textit{RRM} domains are identical in \textit{SUP-26a} and \textit{SUP-26b}. The residue affected by \textit{n1091} is indicated by an arrow.
animals displayed TRA-2::GFP, whereas many cells in sup-26 (gk426) animals expressed TRA-2::GFP. In Western blot analysis of smIs350 hermaphrodites with different genetic backgrounds, we observed significantly increased levels of an ≈50-kDa TRA-2::3xFLAG polypeptide in sup-26 mutant embryos and L4 larvae compared with those in wild-type embryos and L4 larvae (Fig. 3 C and D). This TRA-2::3xFLAG polypeptide is similar in size to the TRA-2 product (TRA-2ic) generated by TRA-3 protease cleavage at the intracellular domain of TRA-2a (14) and to the predicted size of the TRA-2b isoform. It is also consistent with the size of the TRA-2 protein detected in immunoblot analysis using an antibody raised against the TRA-2 intracellular domain (15, 16). In sup-26(gk426); tra-3(e1107) smIs350 animals, we observed one additional high-molecular-weight form of TRA-2::3xFLAG consistent in size with full-length TRA-2 (Fig. 3D), indicating that TRA-2a is indeed cleaved by TRA-3 in C. elegans. In sup-26(gk426); smIs350/+ males, we observed a similar increase in the abundance of the 50-kDa TRA-2::3xFLAG polypeptide when compared with wild-type smIs350/+ males (Fig. 3E), which were mildly feminized due to TRA-2 overexpression from the smIs350 transgene (Table S2). The feminization phenotype of sup-26(gk426); smIs350/+ males was stronger than that of smIs350/+ males, which is consistent with more increased TRA-2 expression in sup-26(gk426); smIs350/+ males. These results suggest that in both males and hermaphrodites SUP-26 represses tra-2 protein expression. Moreover, sup-26 can inhibit translation from both tra-2 transcripts, which are transcribed from different promoters but share the same 3’ UTR (Fig. S1) (8). Real-time quantitative RT-PCR analysis revealed a slight decrease in tra-2 transcripts in sup-26(gk426) mixed-stage animals when compared with wild-type animals, indicating that sup-26 does not inhibit tra-2 transcription or reduce tra-2 mRNA stability (Fig. 3F). Therefore, our results are consistent with the model that sup-26 regulates tra-2 expression by inhibiting tra-2 translation.

SUP-26 Regulates tra-2 Expression Through the TGE Elements. It was previously shown that translation of tra-2 in the germline is...
repressed by elements in its 3′ UTR (20). We thus tested whether sup-26 acts through the tra-2 3′ UTR. We generated GFP reporters that lack the TRA-2 coding sequence but contain the 816-bp tra-2 promoter, the coding region for nucleus-localized GFP (NLS::GFP), and an 848-bp tra-2 3′ UTR (Fig. 4A; Methods). An integrated transgene, smls236 (Ptra-2-NLS::GFP::3′ UTRΔAAA), had stronger GFP expression in sup-26(gk426) animals than in wild-type animals on the basis of the immunoblotting analysis (Fig. 4B) and the analysis of GFP fluorescence intensity (Fig. 4C). Increased NLS::GFP expression in sup-26(gk426) animals was apparent in most tissues and was particularly obvious in the uterus. For example, in smls236 animals, an average of 17% uterine cells had visible NLS::GFP expression, compared with an average of 70% in sup-26 (gk426); smls261 animals (Fig. 4D). In contrast, a similar integrated transgene lacking both 28-bp TGEs, smls261 [Ptra-2-NLS::GFP::3′ UTR(ΔTGE)6], produced similar levels of NLS::GFP expression in sup-26(gk426) and wild-type animals (Fig. 4B and D), suggesting that SUP-26 likely inhibits tra-2 translation through TGEs.

We then tested whether SUP-26 binds directly to the tra-2 3′ UTR in vitro. We found that a purified SUP-26 GST fusion (GST::SUP-26ΔRRM), which contains the SUP-26b RRM domain (amino acids 81–259), formed a complex with a 32P-labeled TGE RNA oligonucleotide, displaying retarded mobility in a gel shift assay (Fig. 4E, lanes 1 and 2). Unlabeled TGE oligonucleotide competed effectively for binding to GST::SUP-26ΔRRM in a concentration-dependent manner, blocking the complex formation (Fig. 4E, lanes 3–6). In contrast, an RNA oligonucleotide with the identical nucleotide composition but a scrambled sequence was much less effective in doing so, showing an approximately ninefold lower binding affinity (Fig. 4E, lanes 7–10). These results suggest that SUP-26 binds specifically to the 3′ UTR of the tra-2 mRNA through the 28-nt TGEs.

Polyadenylate-Binding Protein Associates with SUP-26 in Vivo. To identify factors that may act with SUP-26 to regulate tra-2 translation, we immunoprecipitated SUP-26::GFP from extracts...
of \( P_{\text{sup-26}}:\text{gfp} \) transgenic animals (Methods). SDS-polyacrylamide gel resolution of proteins coprecipitated with SUP-26::GFP revealed the presence of two major protein bands that were not observed in the mock immunoprecipitation (IP) sample (Fig. S4). MALDI-TOF mass spectroscopy analysis determined that the lower band (Fig. S4, band 2) corresponds to SUP-26::GFP and the upper band (Fig. S4, band 1) corresponds to the poly(A)-binding protein PAB-1 (Table S3), which was confirmed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis using LTQ Orbitrap (Fig. S2). To examine whether SUP-26 and PAB-1 directly interact, we performed a GST fusion protein pulldown assay. GST::SUP-26 and GST::SUP-26C [which contains the carboxyl terminal domain of SUP-26a (amino acids 260–357), but not GST], GST::SUP-26RRM 

or GST::SUP-26N [which contains the amino terminal domain of SUP-26a (amino acids 1–80)] specifically pulled down His::PAB-1::FLAG in the presence of RNase A (Fig. 5B). These results suggest that SUP-26 and PAB-1 can directly interact in vitro independently of RNA through the carboxyl terminal domain of SUP-26. Finally, we tested whether \( sup-26 \) may affect the length of the poly(A) tail of \( tra-2 \) mRNA using a PCR-based assay (Fig. S3) (23, 24). We were able to detect relatively short poly(A) tails on \( tra-2 \) mRNAs, similar to what has been reported previously (23, 24), but their lengths were not affected by mutations in \( sup-26 \) (Fig. S3B). In summary, we have identified an RRM-containing protein, SUP-26, that is ubiquitously expressed in \( C. elegans \) somatic cells, binds specifically to TGEs in the 3′ UTR of \( tra-2 \) mRNA, and modulates somatic sex determination by repressing \( tra-2 \) translation. Interestingly, GLD-1, a germ-cell–specific RNA-binding protein that shares no sequence similarity with SUP-26, also binds TGEs in the 3′ UTR in the germline and to promote spermatogenesis (25, 26). It appears that GLD-1 and SUP-26 use different cofactors or mechanisms to repress \( tra-2 \) translation. In the germline, FOG-2, a unique F-box protein and a germ-cell–specific factor, is proposed to act as a bridge to bring GLD-1–bound \( tra-2 \) mRNA into a translational repression complex (26). We find that PAB-1, a poly(A)-binding protein, associates with SUP-26 in vivo and interacts directly with SUP-26 in vitro independently of RNA. The PABPs have been shown to interact with translation initiation factors such as eIF4G to form a circular mRNA structure that facilitates active translation (1, 29). Their binding to the poly(A) sequences could also prevent deadenylation and thus stabilize mRNAs (1, 29), although loss of \( sup-26 \) does not appear to affect the lengths of \( tra-2 \) poly(A) tails. It seems more likely that the association of SUP-26 with PAB-1 at the \( tra-2 \) 3′ UTR interferes with PAB-1′s function in stimulating \( tra-2 \) translation. If so, this would represent a different TGE-mediated translational repression mechanism from the one used in the germline and perhaps is similar to that used by inhibitory PABP-interacting proteins, which inhibit translation by antagonizing the translation-stimulating activity of PABPs in mammalian cells (30).

**Methods**

**Strains.** Strains were maintained using standard procedures. Transgenic strains were generated by microinjection (31). Integration of extrachromosomal trangene arrays was performed by \( \gamma \)-irradiation method (32). Mutations and integrated arrays used in this study were as follows: LGIII—\( sup-26(ku1834, gk426, gk403, n1091, ct49) \); LGV—\( tra-2(\text{e}1107) \); smIs350 (\( P_{\text{tra-2}}:\text{3xFLAG} \)), smIs261 (\( P_{\text{tra-2}}:\text{NLS::GFP::3′UTR(TGE)}_{\text{tra-2}} \)), LGV—\( unc-76(e911), \) her-1(\( n695 \)), and LGG—\( smIs26 \) (\( P_{\text{tra-2}}:\text{NLS::GFP::3′UTR(TGE)}_{\text{tra-2}} \)) and smIs259 (\( P_{\text{tra-2}}:\text{sup-26::gfp} \)). The chromosomal location of \( smIs380 \) (\( P_{\text{tra-2}}:\text{gfp} \)) has not been determined.

**Molecular Biology and Transgenic Animals.** Sequences of all primers used in this study are listed in Table S4. Fosmids were injected into \( sup-26(n1091) \); her-1(\( n695 \)) animals at 10 ng/\( \mu L \) with pRF4 as a co-injection marker (50 ng/\( \mu L \)). \( P_{\text{sup-26}}:\text{gfp} \) was constructed by PCR amplification of the 4-\( kb \) \( sup-26 \) fragment using the primers \( sup-26 \text{pro} \) and \( sup-26 \text{cas} \) and by subcloning the PCR fragment into a modified pdPD117.01 vector using the standard Gateway cloning technique. \( P_{\text{sup-26}}:\text{gfp} \) was injected at 5 ng/\( \mu L \) with pRF4 (50 ng/\( \mu L \)).

The \( sup-26 \) cDNAs were amplified from a cDNA library prepared from mixed-stage wild-type animals using primers complementary to the predicted 5′ and 3′ ends of the \( sup-26 \)-coding sequence. The amplified cDNA fragments were cloned into the Gateway vector pDONR221. Of 20 \( sup-26 \) clones analyzed by restriction enzyme digestion and DNA sequencing, 17 were \( sup-26 \) and 3 were \( sup-26b \).

The \( P_{\text{tra-2}}:\text{NLS::GFP::3′UTR(TGE)}_{\text{tra-2}} \) reporter was generated by inserting an 816-bp \( tra-2 \) operon promoter fragment (Xbal-Xmal) and an 848-bp \( tra-2 \) 3′ UTR fragment (EcoRlSpeI) into pdPD122.66. \( P_{\text{tra-2}}:\text{NLS::GFP::3′UTR(TGE)}_{\text{tra-2}} \) lacks two TGEs, was generated by site-directed mutagenesis. These plasmids were injected individually into \( unc-76(e911) \) animals at 50 ng/\( \mu L \) with p76-16b (an \( unc-76 \) rescuing plasmid) at 25 ng/\( \mu L \). To generate \( P_{\text{tra-2}}:\text{3xFLAG} \), a 12,850-bp genomic fragment containing an 816-bp promoter upstream of \( pp-1 \), the coding region of \( pp-1 \), and the \( tra-2 \)-coding region were fused to three tandem copies of theFLAG tag (DYKHDDHGYKDHID- DYKDDDDK). The 848-bp \( tra-2 \) 3′ UTR was then fused to the 3′ end of the 3xFLAG tag. \( P_{\text{tra-2}}:\text{gfp} \) was made by replacing the 3xFLAG epitope sequence of \( P_{\text{tra-2}}:\text{3xFLAG} \) with a KpnI-EcoRI gfp fragment from pdPD59.75. \( P_{\text{tra-2}}:\text{3xFLAG} \) or \( P_{\text{tra-2}}:\text{gfp} \) was injected into \( unc-76(e911) \) animals at 25 ng/\( \mu L \) with p76-16b (50 ng/\( \mu L \)).

**Fig. 5.** PAB-1 associates with SUP-26 both in vivo and in vitro. (A) Lysates from \( C. elegans \) animals expressing SUP-26::GFP (smIs259) were prepared as described in Methods, incubated with a mouse anti-GFP monoclonal antibody (GFP IP) or no antibody (mock IP), precipitated using Protein G Sepharose beads, resolved by 12% SDS/PAGE, and subjected to silver staining. Two major bands not observed in mock IP were excised from the gel, subjected to trypsin digestion, and analyzed by MALDI-TOF mass spectroscopy and LC-MS/MS. The upper band corresponds to PAB-1 and the lower band is SUP-26::GFP. (B) PAB-1 associates with SUP-26 in vitro through the carboxyl-terminal domain of SUP-26 in the presence of RNases. A total of 200 ng of purified GST, GST-SUP-26, GST-SUP-26Δ, GST-SUP-26RM, and GST-SUP-26C were incubated with glutathione Sepharose beads and 100 ng of purified His::PAB-1::FLAG. The bead-bound proteins were resolved by 12% SDS/PAGE and analyzed by immunoblotting with anti-GST and anti-FLAG antibodies, respectively. Asterisks indicate the corresponding GST fusion proteins.
Protein Purification and Gel Mobility Shift Assay. Gel shift assays were performed as described previously (33). Briefly, GST:Sup-26 was purified from the human neuroblastoma cell line SH-SY5Y cells using glutathione-Sepharose beads (GE Healthcare). An RNA oligonucleotide corresponding to the 28 nt 3'UTG element was synthesized (Integrated DNA Technologies) and end-labeled with 32P using polynucleotide kinase (New England Biolabs). For the binding reaction, GST:Sup-26 was incubated at 25 °C with 32P-labeled RNA in a binding buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 100 mM KCl, 0.1 mM DTT, 5% glycerol, and 0.1 mg/ml BSA in the presence or absence of unlabelled RNA oligonucleotide competitors. After a 20-min incubation, the samples were resolved on a 5% nondenaturing polyacrylamide gel at 4 °C. The gel was then dried and exposed to a Phosphorimaging screen (Perkin-Elmer).

Microscopy Imaging. Fluorescence and differential interference contrast (DIC) images were collected at 0.5-μm intervals with an Axioscan 2 microscope (Zeiss) and a cooled CCD camera (PCO SensiCam). Fluorescence images were subjected to deconvolution analysis using the Slidebook 5.0 software program (Intelligent Imaging Innovations).

Mass Spectroscopy Analysis. Mixed-stage animals were harvested from nematode growth media (NGM) agar plates and lysed by sonication (3× 10 s) in a buffer containing 250 mM NaCl, 100 mM Tris-HCl (pH 7.4), 1 mM EDTA, 5 mM Na2EDTA, 0.1% Triton X-100, 1% PMSF, 5 mM DTT, and 1% Pefabloc Complete Protease Inhibitor Mixture. Lysates were confirmed by centrifugation at 14,000 × g for 30 min, precleared with Protein G beads (GE Healthcare), and then incubated with an anti-FLAG antibody and Protein G beads for 2 h at 4 °C with gentle rocking. After four extensive washes with the same buffer, the precipitated samples were resolved in 12% SDS/PAGE and silver-stained. In-gel tryptic digestion of silver-stained proteins and mass spectrometric analysis was carried out as described (SI Methods) (34).


SI Methods

Mass Spectroscopy Analysis. The protein bands of interest excised from silver-stained gels were destained by 1% potassium ferricyanide and 1.6% sodium thiosulfate, subjected to reduction and alkylation by 10 mM DTT/55 mM iodoacetamide in 25 mM NH₄HCO₃, and then in-gel digested with trypsin (20 μg/mL in 25 mM NH₄HCO₃) at 37 °C for 16 h. The tryptic peptides were then analyzed by mass spectrometry. For MALDI-TOF mass spectrometric analysis, tryptic peptides were mixed with α-cyano-4-hydroxycinnamic acid (CHCA) matrix containing 2 fmol internal standards and analyzed on an Ultraflex MALDI-TOF mass spectrometer (Bruker Daltonics). For liquid chromatography-tandem mass spectrometry (LC MS/MS) analysis, each peptide mixture was reconstituted in HPLC buffer A (0.1% formic acid), loaded across a trap column (Zorbax 300SB-C18, 0.3 × 5 mm, Agilent Technologies), and separated on a 10-cm analytical C18 column (inner diameter, 75 μm; New Objective). The peptides were eluted using a linear gradient of 0–95% HPLC buffer B (99.9% acetonitrile containing 0.1% formic acid). The LC apparatus was coupled with a 2D linear ion trap mass spectrometer (LTQ-Orbitrap, Thermo Fisher). The resulting spectra from both mass spectrometers were used in searches of the National Center for Biotechnology Information nr database with taxonomy set on Caenorhabditis elegans (26,209 sequences) assuming the digestion enzyme trypsin. The MASCOT search engine (http://www.matrixscience.com; v.2.2.03 Matrix Science) was used, allowing one missed cleavage site with charge states from 1+ to 3+. MS mass tolerance was set to be 10 ppm for LC-MS/MS (50 ppm for MALDI-TOF MS), and MS/MS tolerance was set to be 0.5 Da for fixed oxidation for carbamidomethyl cysteines and variable oxidation of methionine residues. Protein identification was performed using Mowse scores (P < 0.05) and the MudPIT algorithm of the MASCOT search engine. Peptide identification was performed using the Scaffold 2 search engine (v.2.06.01; http://www.proteomesoftware.com).

Fig. S1. tra-2 transgenes. (A) Schematic of the P_tra-2::gfp and P_tra-2::3xflag translational fusions and the P_tra-2(intron 17)nls::gfp transcriptional reporter. Boxes represent exons, lines indicate introns or promoter, and the triangle represents the 3′ untranslated region. The tra-2b transcript is also shown. The ability of the translational fusions to rescue the defects of the tra-2(lf) mutants is indicated. The tra-2a transcript is trans-spliced to the SL2 splice leader, which is consistent with its being a downstream gene in an operon. The tra-2b transcript is trans-spliced to the SL1 splice leader, indicating that it is not expressed as a part of the operon (1, 2). The gene upstream of tra-2 in the operon is ppp-1. Green box indicates GFP. Red box indicates the 3xFLAG tag. (B) Images of transgenic embryo and larva carrying the P_tra-2(intron 17)nls::gfp transgene, which directs nucleus-localized GFP expression under the control of the 2071-bp intron 17 of the tra-2a isoform.

A PolyA Binding protein family member (PAB-1) [Caenorhabditis elegans]  
(NCBI Reference Sequence: NP_001021711.1)

1 MemVAApaa AvagaAapQP GQQTGGSYY MASLYVGDLH PDDVNESILFE KFAAAGPvLS
61 IvrcrDntrL LslgYaYVNf QQPDAEAM DTMNFELAHG KPRM1MSQR DPAAMRSQAg
121 NliELnLDK FVgkFQPrAQ RNEELGTAQ LOGTVYYKNF GQHYKETGL KLFKFGNIT
181 CcVeMrTVEG SkgFgPaFa NPEEAETAVQ ALHDSTIEGT DLLKlHCRAQ KKSREALEK
241 KkHEQHkAer MQkYQGVYI LKfNTLEYDD ALIKKOFesh GNTSakVMT DENGRSKFGF
301 FvcFeKePeeA TsaVTEmsk MvcsKpLyV A TAOERk Rdra 0laasQaMb AsRmHNgvP
361 GaAMYNTPQ PGPYVYANPM SCNYFlPQGQQ MQVEPGrGWR MUNrYyvPNQ YrMQAGgvYy
421 QNRmGPQnqQ QGGPGPQPOQQ YNVQAQGGVR MQGPRrTPQG rVGGVQVFPQ rPQQQQQRFa
481 PtGKPAPpQP YQAQYRQPQ ViYGGQEpIT SAMLAAAPQ EQQQLlGERI YALEKLEYPG
541 HkDaGKTGm MLIDNSeLI MMLQDSeLfr SKVDEAASVL VSAQKQ

Tryptic peptide GPPQQYNOVAQGGVR

B C. elegans protein R10E4.2a (SUP-26a) [Caenorhabditis elegans]. GenBank: CAA90772.1

1 NNAASAPQQ QQQQQQQQQAP PQQQQQQPQ HrQVHCmHPg SFGFgNSgY
61 GYraAPQrQO HHDDTtPllS TNNLYGBLMP NtNdDlREm CSyGyNiaSt KAIrDkATNN
121 CkGyGFDPE SPqAaaaAVD GLmtEGtQAq MAKoQQcEQd PTNLyIANL PDLTeQMLt
181 ElNKFGMvTS THlRTPPDQ SrgVcFArMD skEKEcVtIS AlNGRDPtM skECpAlLk
241 QADTrGSKH SMNNPEmLRQ MQYPGQVsY QyHyPAAvY QHyDVNLsAS QMcoHIVGgG
301 NpQAnGGGDM YGAhMYQgNA gGGGqQqQqGQ GQGQpQQqQFY NPNNRnxKk FYwIIPNPO

Tryptic peptide FMWiST (oxidation (M))

Tryptic peptide CEVIASNLGGR (carbamidomethyl (C))

Fig. 52. Representative MS/MS spectra from LTQ Orbitrap used to confirm the identification of poly(A)-binding protein PAB-1 and SUP-26 in the SUP-26::GFP-containing immunoprecipitated complexes. Tryptic peptides of protein bands 1 and 2 in SDS/PAGE (Fig. 5A) were analyzed by LC-MS/MS using LTQ Orbitrap. The amino acid sequences of peptides identified by MS/MS analysis and matched to the amino acid sequences of PAB-1 (A) and SUP-26 (B) are underlined. The representative MS/MS spectra of two selected peptides (indicated in red) for each protein are shown below the amino acid sequences. The assignments of the fragmented ions observed to specific amino acid residues were performed using the Scadffold 2 search engine, and the search results are shown below the MS/MS spectra.
**Fig. S3.** Analysis of tra-2 mRNAs using the poly(A) tail length assay (PAT). (A) A schematic of RT-PCR products generated by two different primer pairs. Oligo 1 and oligo 2 correspond to different sequences in the tra-2 3′ UTR, respectively. An asterisk indicates $^{32}$P end-labeling of the two oligos by the polynucleotide kinase. The indicated PCR product sizes represent amplicons with a minimum poly(A) tail length of 12 nucleotides. PCR products larger than 95 bp or 263 bp suggest that the corresponding mRNAs have poly(A) tails longer than 12As. (B) RT-PCR products of tra-2 mRNAs from the indicated genotypes. The PAT assay products were resolved on a 2.5% agarose gel and visualized by autoradiography. Lane 1 contains 100-bp DNA ladders that were end-labeled by $^{32}$P. Lanes 2–4 contain RT-PCR products derived from the oligo 1 and oligo(dT)$_{12}$ primer pair, and lanes 5–7 contain products derived from the oligo 2 and oligo(dT)$_{12}$ primer pair.

**Table S1.** sup-26 mutations and their suppression of the her-1(n695gf) Tra defect

<table>
<thead>
<tr>
<th>Genotype</th>
<th>% CEM surviving*</th>
<th>% HSN surviving*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type (male)†</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Wild type</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>sup-26(n1091)</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>her-1(n695)</td>
<td>86</td>
<td>0</td>
</tr>
<tr>
<td>sup-26(n1091); her-1(n695)</td>
<td>14</td>
<td>96</td>
</tr>
<tr>
<td>sup-26(n1091); her-1(n695); smIs336[Psup-26::gfp]</td>
<td>89</td>
<td>0</td>
</tr>
<tr>
<td>sup-26(gk403); her-1(n695)</td>
<td>24</td>
<td>100</td>
</tr>
<tr>
<td>sup-26(ct49); her-1(n695)</td>
<td>31</td>
<td>60</td>
</tr>
<tr>
<td>sup-26(gk426); her-1(n695)</td>
<td>39</td>
<td>48</td>
</tr>
</tbody>
</table>

*The presence of CEMs and HSNs was determined in young adult hermaphrodites (except noted otherwise) with the aid of the integrated smIs26 transgene that carries both Ppkd-2::gfp and Pph-1::gfp reporters, which direct GFP expression in CEMs and HSNs, respectively. The percentage of surviving CEMs was calculated by dividing the total number of CEMs observed by the presumptive number of CEMs [(the number of CEMs observed)/(4 x the number of animals scored)]. The percentage of surviving HSNs was calculated by dividing the number of animals with at least one HSN by the number of animals scored. For all genotypes, 25 animals were scored.

†The wild-type male strain contained the him-5(e1490) mutation that results in an increased frequency of X chromosome nondisjunction and thus an increased frequency of XO males. CEM, cephalic companion neurons; HSN, hermaphrodite-specific neurons.

**Table S2.** sup-26 mutations enhance feminization of males when TRA-2 is overexpressed

<table>
<thead>
<tr>
<th>Genotype</th>
<th>% CEM surviving</th>
<th>% HSN surviving</th>
<th>% abnormal male tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>sup-26(n1091)</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>sup-26(gk426)</td>
<td>99</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>smIs350/+</td>
<td>91</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>sup-26(gk426); smIs350/+</td>
<td>70</td>
<td>0</td>
<td>40</td>
</tr>
</tbody>
</table>

The presence of CEMs and HSNs was determined in young adult males with the aid of the smIs26 transgene. The percentage of surviving CEMs or HSNs was calculated as described in Table S1. The smIs350 integrated transgene carries Ptra-tra-2::3xflag, which directs expression of the TRA-2::3xFLAG protein from the native tra-2 promoter (Methods). For all genotypes, at least 15 animals were scored.

*All strains contain the him-5(e1490) mutation that results in an increased frequency of X chromosome nondisjunction and thus an increased frequency of XO males. CEM, cephalic companion neurons; HSN, hermaphrodite-specific neurons.
Table S3. MALDI-TOF mass spectroscopic analysis of protein bands detected in the SUP-26::GFP-containing immunoprecipitated complexes

<table>
<thead>
<tr>
<th>Band no./protein name</th>
<th>M.W. (kDa)</th>
<th>Mascot score</th>
<th>Sequence coverage rate (%)</th>
<th>Accession no.</th>
<th>Peptide: start–end</th>
<th>Mass expected</th>
<th>Mass calculated</th>
<th>ΔMass (ppm)</th>
<th>Peptide sequence matched</th>
</tr>
</thead>
</table>
| 1/PAB-1 | 71.945 | 59 | 21% | gi71993203 | 52–62 | 1116.6396 1116.6291 | 7 | K.FSAAGPVLISR.V  
Oxidation (M) |
| 2/TAG-310 (SUP-26) | 39.455 | 76 | 25% | gi17554534 | 50–62 | 1268.5529 1268.5534 | 0 | R.GSGPFGNSNYGYR.Y  
Oxidation (M) |
| 68–86 | 2168.0229 2168.0247 | 4 | R.GDOQHHDSTPLSTNLYIR.G  
Oxidation (M) |
| 87–98 | 1373.6593 1373.6609 | –1 | R.GLMPNTDLLR.E  
Oxidation (M) |
Oxidation (M) |
| 185–192 | 909.4783 909.4742 | 10 | R.GMVSTR.I |
| 185–192 | 925.4676 925.4569 | 9 | R.GMVSTR.I oxidation (M) |
| 213–226 | 1544.7862 1544.7981 | –8 | K.EKCEVIALNGLR.F  
Oxidation (M) |
| 215–226 | 1287.6580 1287.6605 | –2 | K.CEVIALNGLR.F  
Oxidation (M) |
| 233–240 | 839.5134 839.5116 | 2 | K.EGAPALLI.K |
| 248–260 | 1570.7263 1570.7344 | –5 | K.SHSMNPNPEMLQR.M  
Oxidation (M) |
| 250–260 | 1355.6149 1355.6074 | 6 | K.HSMNPNPEMLQR.M  
Oxidation (M) |
| 250–260 | 1371.6174 1371.6023 | 11 | K.HSMNPNPEMLQR.M  
Oxidation (M) |

Tryptic peptide mass fingerprints of protein bands 1 and 2 in the SDS/PAGE (Fig. 5A) were analyzed by MALDI-TOF mass spectrometry. When searched in the National Center for Biotechnology Information nr database with taxonomy set on C. elegans, these observed peptide masses matched to 11 peptides of PAB-1 and to 13 peptides of TAG-310 (SUP-26), with a 21% and a 25% sequence coverage rate, respectively. The corresponding amino acid residue numbers (start to end) of matched peptides in PAB-1 and SUP-26 are denoted. The Mascot score used for protein identification is defined as $-10 \times \log(P)$, where $P$ is the probability that the observed match is a random event. Scores greater than 57 are considered to be significant ($P < 0.05$). ΔMass (in ppm) indicates the difference between expected (expt) mass and calculated (calc) mass, which is defined as $[\text{mass(expt)} - \text{mass(calc)}]/\text{mass(calc)} \times 1,000,000$. M.W., molecular weight.

Table S4. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUP-26proS</td>
<td>5′ GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGCTAGCTCGCAGAGCCTAAAAACATTTC 3′</td>
</tr>
<tr>
<td>SUP-26NS</td>
<td>5′ GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGAACGCATCTTCGGCTCCACAACAAC 3′</td>
</tr>
<tr>
<td>SUP-26CAS</td>
<td>5′ GGGGACCACTTTGTACAAAAAAGCAGGCTTAATGAACGCATCTTCGGCTCCACAACAAC 3′</td>
</tr>
<tr>
<td>TRA-2proS</td>
<td>5′ CGGGATCCCCATGGTGAGGCCT 3′</td>
</tr>
<tr>
<td>TRA-2proAS</td>
<td>5′ AAATCCCGGGTTTTATCTGAAATTTGTTATTTGAACG 3′</td>
</tr>
<tr>
<td>TRA-2 3′S</td>
<td>5′ ATGAATTCAATGTCTGTTTCCTTTTTCAG 3′</td>
</tr>
<tr>
<td>TRA-2 3′AS</td>
<td>5′ ATATACTAGTTAATACCTAAGTACATTTACGTATATAAC 3′</td>
</tr>
<tr>
<td>TRA-2 3′T7 S</td>
<td>5′ TATAACGACTCTATAGGGAATGCTGGTTCTTTCTTTCTTTCTAATTGCCATTTG 3′</td>
</tr>
<tr>
<td>TRA-2 delta TGE S</td>
<td>5′ TGTACAATTTCCATTTCATATCGTCCACTCGAC 3′</td>
</tr>
<tr>
<td>TRA-2 delta TGE AS</td>
<td>5′ GTCGAGTGGACGATATGAAATGGAAATTGTACA 3′</td>
</tr>
<tr>
<td>oligo(dT)12</td>
<td>5′ GCGAGCTTCCGGGGCGCGTTTTTTTTT 3′</td>
</tr>
<tr>
<td>TRA-2 oligo 1</td>
<td>5′ AATGTCTGTTTCCTTTCTCGAATTGGG 3′</td>
</tr>
<tr>
<td>TRA-2 oligo 2</td>
<td>5′ ATCGTCACCTGACCTCAACTTGTAAT 3′</td>
</tr>
</tbody>
</table>