Cis- and *Trans*-Regulatory Mechanisms of Gene Expression in the ASJ Sensory Neuron of *Caenorhabditis elegans*

María González-Barrios,* Juan Carlos Fierro-González,† Eva Krpelanova,‡ José Antonio Mora-Lorca,*,§ José Rafael Pedrajas, Xenia Peñate,*,†† Sebastián Chavez,*,†† Peter Swoboda,† Gert Jansen,‡ and Antonio Miranda-Vizuete*,[1]**

*Instituto de Biomedicina de Sevilla, Hospital Universitario Virgen del Rocío/Consejo Superior de Investigaciones Científicas/Universidad de Sevilla, E-41013 Sevilla, Spain, †Karolinska Institute, Department of Biosciences and Nutrition, S-141 83 Huddinge, Sweden, ‡Department of Cell Biology, Erasmus MC, 3000 CA Rotterdam, The Netherlands, §Departamento de Farmacología, Facultad de Farmacia, Universidad de Sevilla, E-41012 Sevilla, Spain, **Grupo de Bioquímica y Señalización Celular, Departamento de Biología Experimental, Universidad de Jaén, E-23071 Jaén, Spain, and ††Departamento de Genética, Facultad de Biología, Universidad de Sevilla, E-41012 Sevilla, Spain

ORCID IDs: 0000-0003-4871-1621 (J.C.F.-G.); 0000-0001-6416-8572 (P.S.); 0000-0002-7524-171X (G.J.); 0000-0002-6856-5396 (A.M.-V.)

ABSTRACT The identity of a given cell type is determined by the expression of a set of genes sharing common *cis*-regulatory motifs and being regulated by shared transcription factors. Here, we identify *cis* and *trans* regulatory elements that drive gene expression in the bilateral sensory neuron ASJ, located in the head of the nematode *Caenorhabditis elegans*. For this purpose, we have dissected the promoters of the only two genes so far reported to be exclusively expressed in ASJ, *trx-1* and *ssu-1*. We hereby identify the *ASJ motif*, a functional *cis*-regulatory bipartite promoter region composed of two individual 6 bp elements separated by a 3 bp *linker*. The first element is a 6 bp CG-rich sequence that presumably binds the *Sp* family member zinc-finger transcription factor SPTF-1. Interestingly, within the *C. elegans* nervous system SPTF-1 is also found to be expressed only in ASJ neurons where it regulates expression of other genes in these neurons and ASJ cell fate. The second element of the bipartite motif is a 6 bp AT-rich sequence that is predicted to potentially bind a transcription factor of the homeobox family. Together, our findings identify a specific promoter signature and SPTF-1 as a transcription factor that functions as a terminal selector gene to regulate gene expression in *C. elegans* ASJ sensory neurons.

KEYWORDS *Caenorhabditis elegans*; ASJ neuron; promoter; thioredoxin; ASJ motif; SPTF-1 transcription factor

DURING development, spatiotemporal regulation is essential for the generation of the cell-type diversity found in multicellular organisms. The identity of any cell type is driven by the combinatorial interaction of specific *trans*-acting factors, mainly microRNAs and transcription factors (TFs), with *cis*-regulatory elements present in noncoding DNA sequences within or near any given gene. This physical interaction is necessary for the activation and maintenance of gene transcription (Davidson 2006; Levine 2010; Bulger and Groudine 2011). However, the identification of *cis*- and *trans*-acting elements is complex as TFs bind to short and degenerate sequences that appear at high frequency in the genome. Furthermore, TFs sometimes regulate expression by interacting with other DNA-bound transcription factors and/or coregulators, instead of directly binding DNA. In this case, TFs are recruited to genomic locations lacking their binding site, further complicating the process of identification (Junion *et al.* 2012).

In multicellular organisms, individual cell types are defined by the expression of sets of genes that are coregulated by common *cis*-regulatory motifs, termed *terminal differentiation genes* (Morgan 1934; Davidson 2001; Etchberger *et al.* 2007; Flames and Hobert 2009). In turn, the genes that control these terminal differentiation genes usually code for postmitotically expressed transcription factors that directly bind to

doi: 10.1534/genetics.115.176172
Manuscript received October 27, 2014; accepted for publication March 11, 2015; published Early Online March 12, 2015.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.176172/-/DC1.
[1]Corresponding author: Instituto de Biomedicina de Sevilla, Hospital Universitario Virgen del Rocío, E-41013 Sevilla, Spain. E-mail: amiranda-ibis@us.es

the common *cis*-regulatory motifs and are termed *terminal selector genes* (Hobert 2008).

Sequence comparisons of evolutionary divergent organisms have confirmed the conservation of gene-coding sequences throughout millions of years of evolution, as well as pointed out the presence of highly conserved noncoding DNA regions (Bejerano *et al.* 2004; Sandelin *et al.* 2004; Woolfe *et al.* 2005). Similar to mutations in coding sequences, variations in *cis*-regulatory regions can also generate phenotypic diversity (Wray 2007). Therefore, the identification of *cis*- and *trans*-regulatory elements and the study of the molecular mechanisms that control gene expression are essential in developmental biology.

The nematode *Caenorhabditis elegans* combines anatomical simplicity with relatively complex behaviors like sensory responses to touch, smell, taste, and temperature, as well as male mating, social behavior, learning, or memory (Rankin 2002; de Bono 2003). Although a large morphological and evolutionary distance exists between worms and humans, they share many conserved genes and most of the signaling pathways (Lai *et al.* 2000). The *C. elegans* nervous system is one of the most complex tissues in worms and is composed of 302 neurons in the adult hermaphrodite, which represents approximately one-third of its total number of somatic cells (White *et al.* 1986). The ability to monitor and manipulate gene expression in single neurons makes this nematode the ideal model to dissect the *cis*- and *trans*-regulation of neuron-specific gene expression.

Several *cis-trans* modules regulating gene expression in defined neuron types have been identified in *C. elegans*. For example, the CHE-1 terminal selector directly induces the expression of many terminal differentiation genes in the gustatory neurons ASE and controls the expression of various regulatory factors that trigger the asymmetric differentiation of the ASE neurons (Etchberger *et al.* 2007, 2009). Likewise, the LIM homeodomain transcription factor TTX-3 acts as a terminal selector to drive the terminal differentiation program of the serotonergic NSM neurons and the cholinergic interneurons AIY and AIA. The type of differentiation program controlled by TTX-3 in different neuron types is specified by a distinctive combination of additional transcription factors (Wenick and Hobert 2004; Zhang *et al.* 2014). Similar genetic studies on other neuronal types, including AWB (Nokes *et al.* 2009), AWC (Kim *et al.* 2010), and ciliated (Swoboda *et al.* 2000) and dopaminergic (Flames and Hobert 2009; Doitsidou *et al.* 2013) neurons have generated a unique understanding of the *cis* and *trans* mechanisms that initiate single neuronal identities and are responsible for cellular diversity and organism complexity.

C. elegans ASJ neurons are a bilaterally symmetric pair of ciliated sensory neurons located in the amphid, the main chemosensory organ of the nematode. ASJ controls entry and exit from the dauer stage (Bargmann and Horvitz 1991; Schackwitz *et al.* 1996; Cornils *et al.* 2011), an alternative developmental stage whereby the larva survives harsh conditions (Riddle *et al.* 1981). *C. elegans* ASJ neurons also play a role in the regulation of aging (Alcedo and Kenyon 2004), photoreception (Ward *et al.* 2008), and cold habituation (Ohta *et al.* 2014).

To date, only the genes *trx-1* (*thioredoxin-1*) and *ssu-1* (*alcohol sulfotransferase-1*) have been reported to be expressed exclusively in ASJ neurons (Carroll *et al.* 2006; Miranda-Vizuete *et al.* 2006). TRX-1 belongs to the thioredoxin family of redox proteins, playing a role in regulating worm longevity and dauer formation (Miranda-Vizuete *et al.* 2006; Fierro-Gonzalez *et al.* 2011a,b). SSU-1 is a sulfotransferase whose function is required for the volatile anesthetic sensitivity and uncoordinated movement phenotypes of *unc-1* and *unc-24* mutant worms (Carroll *et al.* 2006).

Here we dissect the promoters of *C. elegans trx-1* and *ssu-1* genes and identify *cis*- and *trans*-acting elements that drive expression in this pair of sensory neurons.

Materials and Methods

C. elegans strains and culture conditions

The standard methods used for culturing and maintenance of *C. elegans* were as previously described (Brenner 1974). The strains used in this work are the following: *GJ1186*, *gjIs698 [Psptf-1*::*gfp*; *Pelt-2*::*mCherry]*; *GJ1401*, *gjEx856 [Pgpa-9*::*gfp*; *rol-6(su1006)]*; *GJ2296*, *gjIs698 [Psptf-1*::*gfp*; *Pelt-2*::*mCherry]*; *ofEx205 [Ptrx-1*::*trx-1*::*DsRed*; *lin-15ab(+)]*; *GR1455*, *mgIs40 [Pdaf-28*::*gfp]*; *OE3007*, *lin-15 (n765ts) X*; *ofEx1 [Ptrx-1*::*gfp*; *lin-15ab(+)]*; *VZ126*, *lin-15 (n765ts) X*; *vzEx29 [Pssu-1*::*gfp*; *rol-6(su1006)]*; *VZ250*, *sptf-2 (tm1130) II*; *vzEx87 [Ptrx-1(-300-100)*::*gfp*; *Punc-122*::gfp; *VZ260*, *lin-15(n765ts) X*; *ofEx205 [Ptrx-1*::*trx-1*::*DsRed*; *lin-15ab(+)]*; *vzEx95 [Ptrx-1[tandem 3x(-200, -173)]*::*gfp*; *rol-6(su1006)]*; *VZ289*, *rhIs13 [Punc-119*::*unc-119*::*gfp*; *dpy-20(+)] nre-1(hd20) lin-15b(hd126) X*; *vzEx80 [Ptrx-1 (-342,-84)*::*gfp*; *rol-6(su1006)]*; *VZ382*, *rrf-3(pk1426) II*; *vzEx81 [Pssu-1(-330,0)*::*gfp*; *rol-6(su1006)]*; *VZ383*, *tax-2 (p691) I*; *rrf-3(pk1426) II*; *ofIs1 [Ptrx-1*::*trx-1*::*gfp*; *lin-15ab(+)] IV*; *VZ489*, *sptf-2(tm1130) II*; *vzEx81 [Pssu-1(-330,0)*::*gfp*; *rol-6(su1006)]*; *VZ504*, *sptf-3(gu85) I*; *ofEx1052 [Ptrx-1*:: *trx-1*::*gfp*; *Punc-122*::*DsRed]*; *VZ509*, *sptf-3(gu85) I*; *vzEx81 [Pssu-1(-330,0)*::*gfp*; *rol-6(su1006)]*; *VZ577*, *sptf-1(tm784) I*; *tax-4(p678) III*; *ofIs1 [Ptrx-1*::*trx-1*::*gfp*; *lin-15ab(+)] IV*; *gjEx867 [Pelt-2*::*mCherry*; *sptf-1(+)]*; *VZ578*, *sptf-1(tm784) I*; *gjEx867 [Pelt-2*::*mCherry*; *sptf-1(+)]*; *mgIs40 [Pdaf-28*::*gfp]*; *VZ580*, *tax-2(p691) I*; *cog-1(sy607)/mnC1 dpy-10(e128) unc-52 (e444) II*; *ofIs1 [Ptrx-1*::*trx-1*::*gfp*; *lin-15ab(+)] IV*; *VZ582*, *sptf-1(tm784) I*; *gjEx867 [Pelt-2*::*mCherry*; *sptf-1(+)]*; *vzEx81 [Pssu-1(-330,0)*::*gfp*; *rol-6 (su1006)]*; *VZ583*, *sptf-1(tm784) I*; *gjEx867 [Pelt-2*::*mCherry*; *sptf-1(+)]*; *gjEx856 [Pgpa-9*::*gfp*; *rol-6(su1006)]*; and *VZ584*, *rrf-3(pk1426) II*; *gjIs698 [Psptf-1*:: *gfp*; *Pelt-2*::*mCherry]*.

GFP expression constructs and transgenesis

Ptrx-1(1kb)::*gfp* (Miranda-Vizuete *et al.* 2006) and *Pssu-1 (0.5kb)*::*gfp* plasmids (Carroll *et al.* 2006) were used as the template for generating promoter deletion GFP reporters

as PCR linear fragments to maximize expression efficiency (Etchberger and Hobert 2008). These constructs contained different promoter fragments fused to the *gfp*-coding sequence and the 3′ UTR region of the *unc-54* gene. Mutagenesis reactions (deletions and substitutions) were performed using the QuickChange II XL Site Directed Mutagenesis Kit (Stratagene) and confirmed by sequencing. The 3× tandem-repeat fragment was synthesized *de novo* as two complementary 84 bp oligonucleotides, hybridized, and cloned into the *Hin*dIII and *Bam*HI sites of the *pPD95.77* vector and subsequently used as the basis for generating the corresponding linear fragment by PCR. The *Ptrx-1*::*trx-1*::*DsRed* construct was made by cloning the *trx-1* promoter (1 kb) into the *Hin*dIII and *Bam*HI sites of the pCJ102 plasmid (Winkelbauer *et al.* 2005). Primer sequences used for *trx-1* and *ssu-1* constructs can be provided upon request. The *Psptf-1*::*gfp* construct was generated by cloning a PCR fragment generated with primers forward AGCATGCTTGGGTAAACTGCTTGGGAAAGTGG and reverse AGGATCCGTGATTCTGGTTTTGCTGGCTCTGC into the *Sph*I and *Bam*HI sites upstream of the *gfp*-coding sequence of the vector *pPD95.79*. This construct includes an ∼3.5-kb upstream sequence and fuses the *sptf-1* coding sequence at the end of the second exon to GFP. Transgenic worms were generated by DNA microinjection as previously described (Mello *et al.* 1991). The pRF4 [*rol-6(su1006)*] plasmid was used as injection marker at 50 ng/μl. All fragments were injected into *N2* wild-type animals at varying concentrations ranging from 5 to 50 ng/μl.

Microscopy

GFP transgenic animals were mounted in a 5-μl drop of 10 mM levamisole (Sigma, St. Louis) on a 3% agarose pad covered with a 24- × 24-mm coverslip and scored on an Olympus BX61 fluorescence microscope. For GFP expression analysis in ASJ neurons, at least 50 worms per line/strain were analyzed under a ×20 objective. Cells expressing *Psptf-1*::*gfp* were identified based on their position within the animals by combining fluorescence and DIC microscopy using a Zeiss AxioImager Z1 based on colocalization with DsRed expressed in ASJ neurons from the *trx-1* promoter and based on DiI fluorescent dye filling visualized using a Nikon Ti Eclipse microscope equipped with a Yokogawa CSU-X1 spinning disk unit, 491- and 561-nm lasers, and beam splitters.

Fluorescent dye-filling assays

Fluorescent dye filling was performed using 0.1 mg/ml DiI (Molecular Probes) as previously described (Perkins *et al.* 1986).

RNA interference

Feeding RNA interference (RNAi) was performed as previously described (Timmons *et al.* 2001). RNAi screening was performed in the *VZ289*, *VZ382*, and *VZ383* strains.

Bioinformatics and database analysis

The initial motif-finding analysis in 70 ASJ-expressed genes was performed with the MEME tool (http://meme.sdsc.edu/meme/intro.html). Different lengths (up to 1 kb) of gene upstream regulatory sequences were analyzed. All sequence alignments were performed with the EMBOSS Water Nucleotide Alignment tool (http://www.ebi.ac.uk/Tools/psa/emboss_water/nucleotide.html). The ASJ motif logo shown in Figure 4F was generated with WebLogo (http://weblogo.threeplusone.com). For the identification of transcription factor(s) potentially binding to the ASJ motif, we used the following prediction programs and database resources: Promo 3.0 (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3); Alibaba 2.1 (http://labmom.com/link/alibaba_2_1_tf_binding_prediction); MatInspector (http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html); and Jaspar (http://jaspar.genereg.net/).

Results and Discussion

trx-1 gene promoter analysis identifies a bipartite motif necessary for gene expression in ASJ neurons

To date, ∼70 genes have been reported to be expressed in ASJ neurons (http://www.wormbase.org), whereby only the genes *trx-1* and *ssu-1* have been found to have an expression pattern restricted to this bilateral pair of neurons (Carroll *et al.* 2006; Miranda-Vizuete *et al.* 2006). To determine whether genes expressed in ASJ share a common regulatory motif, we first followed a bioinformatics approach. Alignment of the 70 promoters (each consisting of 1 kb upstream of the ATG codon of the respective ASJ-expressed gene) failed to identify any shared motif owing to the large sequence analyzed (data not shown). Next, we experimentally dissected the 1 kb upstream region of the *trx-1* gene, which is sufficient to drive expression in ASJ neurons (Miranda-Vizuete *et al.* 2006) (Figure 1A). *trx-1* promoter fragments of decreasing lengths were fused to the GFP-coding sequence, and GFP expression in ASJ neurons of transgenic animals was scored. As shown in Figure 1B, we identified a minimal region of 116 bp (positioned at −200 to −84 bp upstream of the *trx-1a* isoform ATG codon) that drives GFP expression in ASJ, thus containing the putative *cis*-regulatory motif responsible for ASJ expression (called "*ASJ motif*" hereafter).

We then made 25 to 30 bp nonoverlapping deletions covering the 116 bp minimal region previously identified. As shown in Figure 1C, deletion of the region −200 to −173 bp upstream of the *trx-1a* ATG codon fully abolished expression in ASJ, whereas deletion of any of the other sequences within the minimal 116 bp region did not affect GFP expression in these neurons. These results indicate that the ASJ motif is located within the region from −200 to −173 bp of the *trx-1a* promoter.

To elucidate if this 28 bp sequence is sufficient to drive expression in ASJ neurons, we generated a 3×28 bp tandem repeat in an otherwise promoter-less GFP reporter. Transgenic embryos carrying this construct showed expression in a pair of head neurons with dendrites extending to the tip of the head, consistent with ASJ morphology and position (Figure 1D). However, once the embryo hatched, the

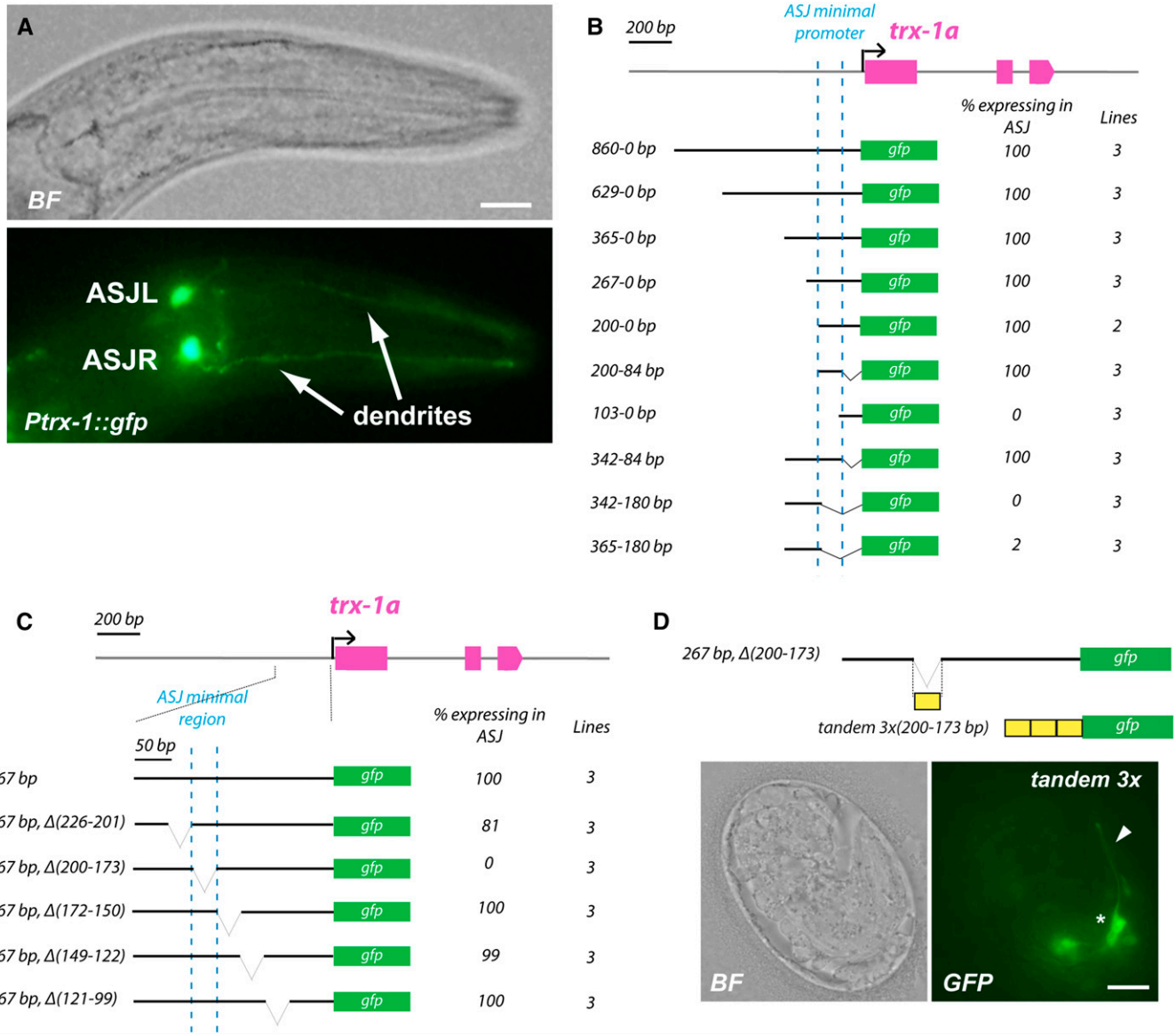

Figure 1 Identification of the minimal *trx-1* promoter region necessary for gene expression in ASJ neurons. (A) Bright-field (BF, top) and fluorescence (bottom) images of *Ptrx-1(1kb)::gfp* transgenic animals. ASJR (ASJ right) and ASJL (ASJ left) neurons are indicated. When expressed from the *trx-1* promoter, GFP is found in the entire cell: in cilia, dendrites, cell bodies, and axons. Bar, 20 μm. (B) Identification of the minimal *trx-1* promoter region that drives gene expression in ASJ neurons. Decreasing length fragments of the *trx-1* promoter fused to the GFP-coding sequence have identified an ASJ minimal promoter (dotted blue line) located at position −200 to −84 upstream of the *trx-1a* ATG codon. (C) Serial deletions within the *trx-1* minimal promoter region previously identified reveal the sequence −200 to −173 (dotted blue line) upstream of the *trx-1a* ATG codon as necessary for gene expression in ASJ neurons. For quantification, at least 50 animals per line were analyzed. (D) Transgenic animals expressing a 3× tandem construct of the sequence −200 to −173 (necessary for gene expression in ASJ) fused to GFP show fluorescence in head neurons with morphology and position consistent with being ASJ. The asterisk indicates the ASJ cell body, and the arrowhead points to the ASJ dendrite. Bar, 5 μm.

number of animals with GFP-fluorescent ASJ neurons decreased dramatically. We could eventually identify very few L1 larvae with expression in ASJ, but none in L2–L4 larvae or adult animals. This is most likely due to the requirement of additional *cis*-regulatory regions in the surrounding promoter area necessary for post-embryonic maintenance of *trx-1* expression, as has been reported for other genes with neuronal expression (Bertrand and Hobert 2009).

Finally, to unequivocally identify the *cis*-regulatory motif that controls *trx-1* expression in *C. elegans* ASJ neurons, we performed a scanning substitution mutagenesis of the 28 bp sequence described above. As shown in Figure 2A, we identified a bipartite motif composed of two 6 bp elements (CAACCC and AATTAA), separated by a 3 bp linker, being absolutely required for ASJ expression. Hence, mutation of either of the two individual 6 bp elements resulted in the abolishment of GFP expression in ASJ neurons, while mutation of the 3 bp linker did not disrupt GFP expression (Figure 2A). Interestingly, a similar bipartite *cis*-regulatory motif has been described to regulate specific expression of genes in the

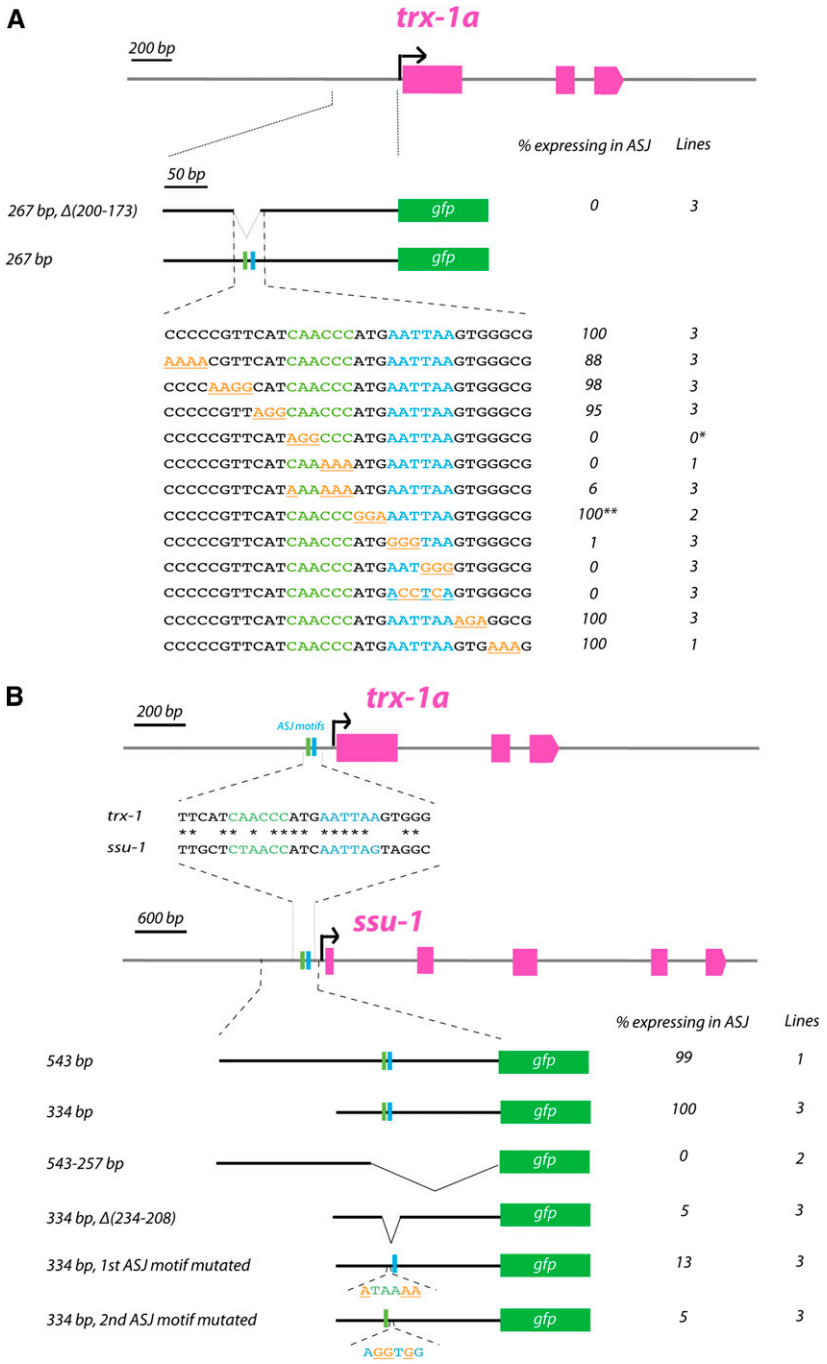

Figure 2 Identification of the ASJ motif in promoters of the *trx-1* and *ssu-1* genes. (A) Scanning substitution mutagenesis (underlined in orange) within the sequence necessary for *trx-1* expression in ASJ (−200 to −173 upstream of the *trx-1a* ATG codon) identifies a bipartite ASJ motif, composed of two individual 6 bp elements (in green and blue, respectively) separated by a 3 bp linker. For quantification, at least 50 animals per line have been analyzed. A single asterisk (*) indicates that we have been unable to get stably transmitting lines from this construct; however, 10 F_1 transgenic animals were scored and no GFP expression in ASJ was observed. A double asterisk (**) indicates that these transgenic animals show several fluorescent head neurons including ASJ. (B) Comparison of the *trx-1* ASJ motif and its homologous region in the *ssu-1* promoter. The diagram below shows that the deletion or the mutation of the individual 6 bp elements of the ASJ motif in the *ssu-1* promoter robustly decreases gene expression in ASJ, indicating that the ASJ motif is also functional in *ssu-1*. For quantification, at least 50 animals per line have been analyzed.

AWB olfactory sensory neurons in *C. elegans* (Nokes *et al.* 2009).

Together, these results indicate that the bipartite motif here identified is necessary and sufficient for *trx-1* expression in ASJ neurons, although additional regulatory sequences are necessary for the (long-term) maintenance of expression in larvae and adult animals.

ssu-1 gene promoter analysis confirms the identity of the ASJ motif

To confirm that the identified ASJ motif was not specific to the *trx-1* gene but conserved in another gene expressed in ASJ neurons, we performed a similar analysis of the *ssu-1* promoter. *ssu-1* together with *trx-1* are the only two genes exclusively expressed in *C. elegans* ASJ neurons (Carroll *et al.* 2006; Miranda-Vizuete *et al.* 2006). A 500 bp promoter region upstream of the *ssu-1* ATG codon has been shown to contain the *cis*-regulatory sequences required for ASJ expression (Carroll *et al.* 2006) (Figure 2B). Using sequential promoter deletions, we found a minimal region of 334 bp upstream of the ATG codon in the *ssu-1* promoter that drives expression in ASJ neurons (Figure 2B). Sequence comparison of this 334 bp region of the *ssu-1* promoter with *trx-1* promoter sequences revealed an *ssu-1* sequence stretch with

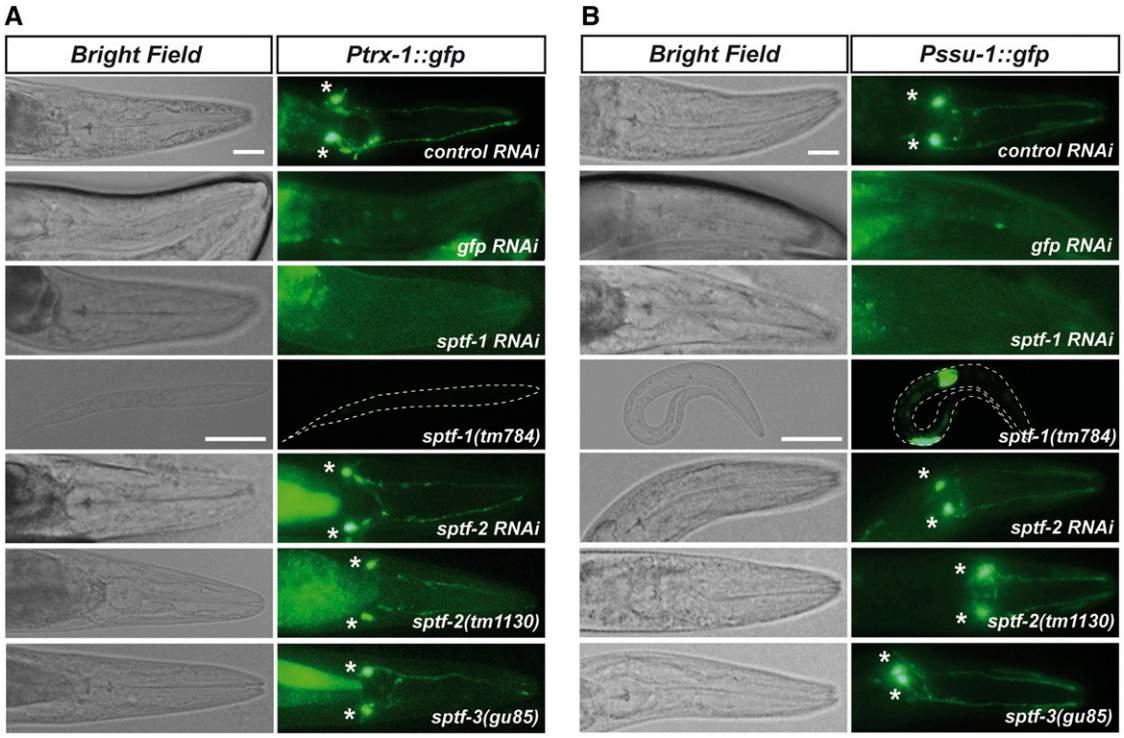

Figure 3 The zinc-finger transcription factor SPTF-1 regulates gene expression in ASJ neurons. (A and B) *sptf-1* mutation or RNAi downregulation abolishes *trx-1* (A) or *ssu-1* (B) promoter-dependent GFP expression in ASJ neurons. RNAi and/or mutant analysis of the other members of the *Sp* family, *sptf-2* and *sptf-3*, demonstrate that these transcription factors are not involved in the regulation of either *trx-1* or *ssu-1* expression in ASJ. Asterisks indicate the ASJ cell bodies. The intestinal labeling of transgenic animals is nonspecific fluorescence due to the *unc-54* 3′ UTR present in the injected constructs (Boulin *et al.* 2006). Bar, 20 μm.

substantial identity to the 28 bp minimal region identified in the *trx-1* promoter containing the ASJ motif. This sequence stretch resides at −234 to −208 bp upstream of the *ssu-1* ATG codon. Indeed, deletion of this region in the *ssu-1* promoter or mutation of each individual element of the bipartite ASJ motif disrupted expression in ASJ (Figure 2B).

Collectively, these results suggest that the bipartite *cis*-regulatory elements identified in both the *trx-1* and *ssu-1* promoters define a novel ASJ motif responsible for gene expression in ASJ neurons in *C. elegans*.

RNAi screening identifies SPTF-1 as a regulator of ASJ expression

The bipartite ASJ motif found in the *trx-1* and *ssu-1* promoters is composed, on the one hand, of a sequence that resembles a CACCC-box and, on the other hand, of an AATTAA-related sequence (Figure 2, A and B). To identify the TFs binding to the ASJ motif, we first used the ASJ motif as bait in several different TF-binding prediction programs and databases available online (*Materials and Methods*). The strongest hits identified in these searches were members of the homeodomain and zinc-finger TF families. This is consistent with the fact that the AATTAA sequence resembles the core consensus sequence recognized by the majority of homeobox TFs (Svingen and Tonissen 2006) while many zinc-finger TFs are known to preferentially bind to GC-rich sites (Kaczynski *et al.* 2003).

Next, using *Ptrx-1*::*gfp* and *Pssu-1*::*gfp* reporter strains, we performed a feeding RNAi screen using the Ahringer TF RNAi sublibrary supplemented with individual homeodomain and zinc-finger TFs not represented in the Ahringer TF sublibrary but that appeared in the bioinformatics search described above (403 RNAi clones, Supporting Information, Table S1). Together, this combined sublibrary represents ∼40% of all predicted transcription factors in the nematode (Reece-Hoyes *et al.* 2005; Shaye and Greenwald 2011). In this screen we identified SPTF-1 (*Specificity Protein Transcription Factor 1*) as the only ASJ motif binding candidate, since *sptf-1* RNAi fully abolished *Ptrx-1*::*gfp* and *Pssu-1*::*gfp* expression in ASJ neurons (Figure 3, A and B). Of note, ASJ labeling is abolished not only in adult animals but also in embryos and larvae, indicating that *sptf-1* deficiency impairs ASJ initiation of gene expression, rather than ASJ gene expression maintenance.

sptf-1 encodes a member of the *Sp* family TF that is composed of multiple zinc-finger-containing proteins that are important components of the eukaryotic cellular transcriptional machinery. *Sp* family TFs bind to GC-rich sites as the GC-box and CACCC-box (Suske 1999; Kaczynski *et al.* 2003). Interestingly, human Sp1 has been found to bind the human *TRX1* promoter and this binding is, in turn, enhanced in a redox-dependent manner by Trx-1 (Bloomfield *et al.* 2003), suggesting that regulation of thioredoxin expression is

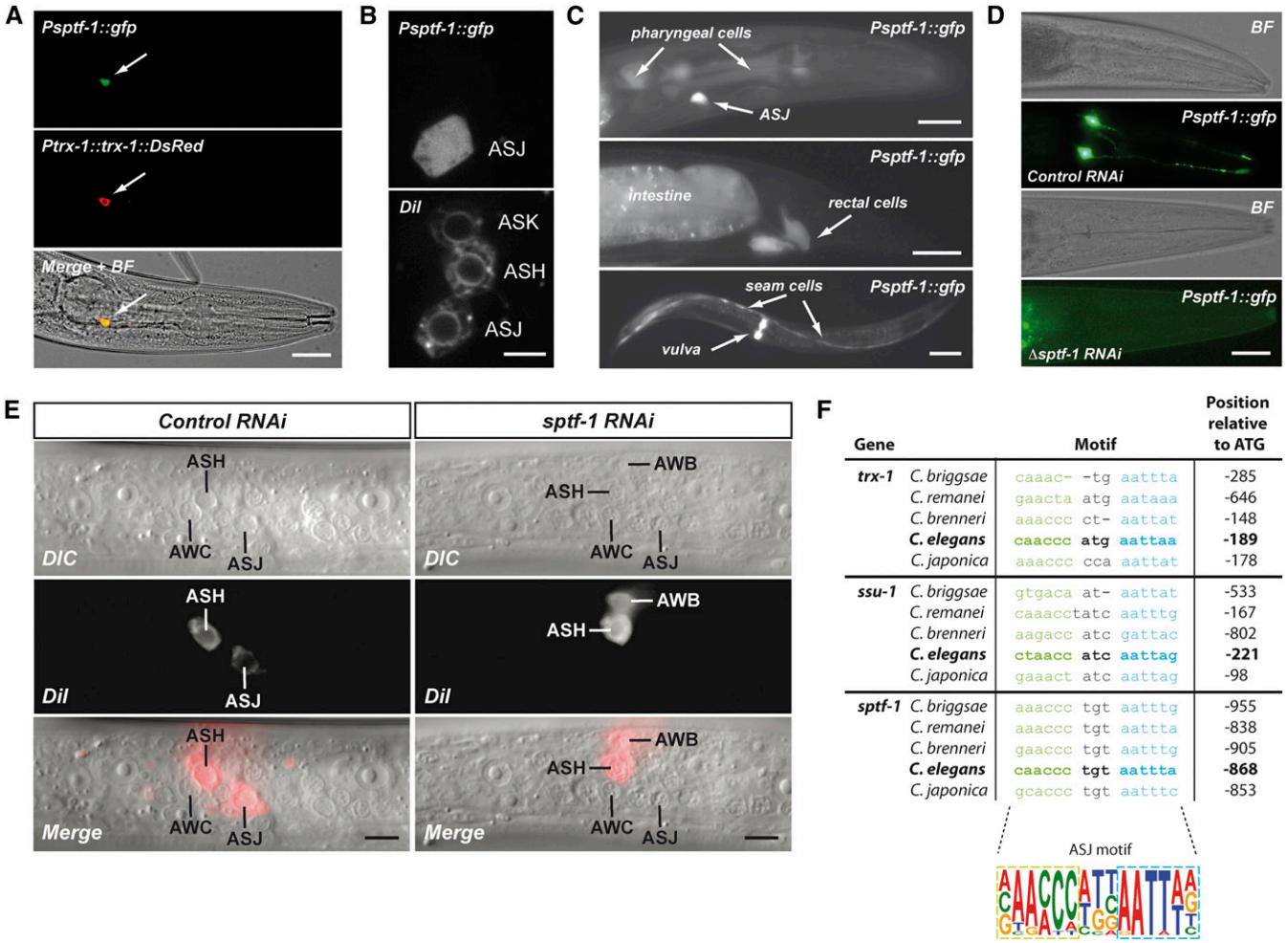

Figure 4 SPTF-1 function in ASJ and conservation of the ASJ motif among *C. elegans* species. (A) Transgenic animals expressing GFP under control of the *sptf-1* promoter and TRX-1::DsRed under control of the *trx-1* promoter demonstrate ASJ colocalization (yellow color in the merged + bright-field image). Bar, 20 μm. (B, top) GFP expressed under control of the *sptf-1* promoter is found exclusively in ASJ neurons. (Bottom) ASJ, ASH, and ASK neurons are stained with the fluorescent DiI dye. Bar, 5 μm. (C) GFP expressed under control of the *sptf-1* promoter is found in other cell types such as pharyngeal cells, rectal cells, intestine, seam cells, and vulval cells. Bar, 20 μm (top and middle) and 100 μm (bottom). (D) RNAi downregulation of the endogenous *sptf-1* gene in a *rrf-3(pk1426)* sensitized background demonstrates that SPTF-1 is autoregulated. Bar, 20 μm. (E) *sptf-1* RNAi downregulation does not interfere with ASJ neurogenesis while it impacts the expression of genes required for DiI uptake. Bar, 10 μm. (F) Sequence alignment of the predicted ASJ motif found in the promoters of the *trx-1*, *ssu-1* and *sptf-1* genes of several *Caenorhabditis* species define the ASJ motif, which is composed of two well-conserved 6 bp elements separated by a 3 bp more divergent linker sequence.

evolutionarily conserved. Similarly, mammalian orthologs of *C. elegans* SSU-1 sulfotransferase have also been reported to be regulated by Sp1 (Hempel *et al.* 2004; Lee *et al.* 2005).

To verify the role of *sptf-1* in regulating ASJ-specific gene expression, we used *sptf-1(tm784)* mutants carrying a deletion that removes the *sptf-1* third exon and causes an L1 larval arrest phenotype. As shown in Figure 3, A and B, *sptf-1(tm784)* arrested L1 larvae carrying the *Ptrx-1*::*gfp* or *Pssu-1*::*gfp* transgenes are devoid of ASJ fluorescence. Notably, a small number of *sptf-1* L1 larvae displayed very weak GFP labeling in the ASJ neurons (Figure S1A), most likely due to the presence of maternal SPTF-1, similar to what has been reported for other members of the *Sp* family of TFs in both *C. elegans* (Ulm *et al.* 2011) and *Drosophila* (Wimmer *et al.* 1996). In other cases, arrested L1 larvae displayed GFP labeling in several neurons including ASJ, suggesting that *sptf-1* deficiency causes an anomalous ectopic expression of the transgenes (Figure S1, B and C).

Importantly, the role of SPTF-1 in controlling gene expression in the ASJ neurons is not shared with other members of the *Sp* family, as neither *sptf-2* RNAi (*sptf-3* RNAi causes a highly penetrant embryonic lethal phenotype that precluded any analysis) nor *sptf-2(tm1130)* and *sptf-3(gu85)* mutants abolished *Ptrx-1*::*gfp* or *Pssu-1*::*gfp* expression in ASJ (Figure 3, A and B). The *sptf-2(tm1130)* mutant carries a deletion that results in a shorter protein lacking part of the zinc-finger domain and therefore presumably is a mutant deficient in DNA-binding activity. Similarly, (*sptf-3gu85*) is a hypomorphic mutant that lacks DNA-binding activity (Ulm *et al.* 2011). Therefore, we conclude that of three

Table 1 Identification of putative ASJ-motif sequences in the promoters of all *C. elegans* genes reported to be expressed in ASJ neurons

Gene	Promoter region analyzed (bp)	Motif	Strand	Position relative to ATG	Position *P*-value	*E*-value
trx-1	200	CAACCCATGAATTAA	+	−189	9.1*E*-08	0.0013
ssu-1	334	CTAACCATCAATTAG	+	−221	1.3*E*-06	0.056
sptf-1	1003	CAACCCTGTAATTTA	+	−868	8.5*E*-07	0.12
dyf-3	650	CAAATCTGGAATTTT	−	−81	4.6*E*-06	0.41
cog-1	2000	CAAATCGGCAATTTG	+	−1487	4.0*E*-06	0.48
		AAAACCGGCAATTTG	+	−1431	5.9*E*-05	
		CAAACCGGCAATTTG	+	−1387	1.7*E*-06	
daf-25	2000	CAAACCGGCAATTTG	−	−1172	1.7*E*-06	0.48
		CAAATCGGTAATTGC	+	−986	3.2*E*-05	
osm-9	1600	CAACTCTTTAATTTT	+	−37	4.0*E*-06	0.89
mks-5	3946	CAAACCGGCAAATTG	−	−3711	4.5*E*-05	0.94
		CAAACCGGCAATTTG	−	−3667	1.7*E*-06	
		CAAATCGGCAATTTG	+	−3379	4.0*E*-06	
		CAACACGGGAATTGT	−	−1224	4.9*E*-05	
C39D10.5	1005	GAAACCGTTAATTAA	−	−240	7.4*E*-06	1
srh-10	1005	CAACCCGCCATTTAT	−	−861	1.7*E*-05	2.3
ceh-13	1051	CGAACCCCCAATCGA	−	−866	1.8*E*-05	2.6
R102.2	700	CTAAGCTTCAATTAA	+	−123	4.3*E*-05	4
odr-8	1005	CAACTCATTGATTGA	+	−483	3.6*E*-05	4.8
ttx-7	524	CAAATCACCATTTGT	−	−436	7.7*E*-05	5.3
ins-6	1714	AAACCCAGCAATTGA	−	−229	2.5*E*-05	5.7
K07C11.10	2289	TGACCCTCCAATTGG	+	−1958	5.9*E*-05	5.8
		CAATTCGCTAATTAG	+	−49	1.9*E*-05	
age1	900	CAACTCCTCAATTCT	+	−494	6.3*E*-05	7.4
ins-4	5716	CTTATCCCTAATTGT	−	−4813	9.5*E*-05	8
		CAACACAGGAATTAA	−	−371	1.1*E*-05	
ins-3	5728	CAACACAGGAATTAA	+	−5314	1.1*E*-05	8
		CTTATCCCTAATTGT	+	−871	9.5*E*-05	
tbx-2	1006	CGAATCAGTAATAAG	+	−790	6.3*E*-05	8.2
tub-1	2800	GAAATCATCAATTAT	−	−2566	8.3*E*-05	10
		CGAACCTGGAATCAA	−	−2468	2.9*E*-05	
		TAAATCTCTAATTAT	+	−1062	6.9*E*-05	
pnc-1	1001	AAAACCTGCAATTTT	+	−340	8.5*E*-05	11
K10D6.2	3450	AAAACCCTTAATTGT	+	−3161	3.5*E*-05	11
		CTAACCTCGATTTTA	−	−2867	5.3*E*-05	
		CGAACTTCTAATTAG	+	−1379	2.5*E*-05	
		CAAATCGTTCATTTG	+	−931	7.7*E*-05	
dyf-2	3000	ATAATCCTGAATTAA	+	−1920	8.9*E*-05	11
		CAAACCGGGAATGAT	+	−235	2.9*E*-05	
unc-103	5000	CAAACAATGAATTGA	+	−4855	2.0*E*-05	13
sra-39	2877	CAAATTGTGAATTAA	+	−624	5.6*E*-05	19
nphp-1	4820	GGAATCACTAATTAA	−	−317	3.6*E*-05	21
gpa-10	3000	CGAACTCCGAATTTG	+	−1739	5.9*E*-05	21
ins-1	4200	CAAATATCTAATTAT	+	−2561	6.9*E*-05	23
		CAAAACTGCAATTAT	−	−1543	4.7*E*-05	
		CTAATAATCAATTAA	−	−1266	8.3*E*-05	
add1	3100	CAACTCATTATTTGT	+	−2288	6.9*E*-05	24
nhr-67	5000	CAAATCGACAATTAT	+	−2621	4.7*E*-05	26
		CTACCACTCAATTGA	+	−331	5.3*E*-05	
dyf-11	4101	CAAATTCCCAATTTG	+	−2825	9.5*E*-05	36
		CAAATATGCAATTGA	+	−2597	9.5*E*-05	
		CAATTCTTTAATTTT	−	−274	8.9*E*-05	
wrk-1	4642	CAAATCATTTATTTT	+	−4469	8.3*E*-05	38
gcy-27	1020			0		13
K07C5.9	425			0		14
arrd-4	1000			0		15
gpa-1	1500			0		19
npr-15	1010			0		20
egap-5	1153			0		24
daf-11	1020			0		25
gpa-3	2000			0		26
cng-2	2362			0		28

(continued)

Table 1, *continued*

Gene	Promoter region analyzed (bp)	Motif	Strand	Position relative to ATG	Position *P*-value	*E*-value
ins-9	1100			0		29
kin-29	830			0		34
c34d4.1	1006			0		34
odr-4	1026			0		35
gpa-9	3000			0		35
egl-43	1734			0		37
tax-2	1982			0		40
ncam-1	2069			0		41
clhm-1	3000			0		42
nlp-3	1024			0		42
npr-5	1010			0		42
nphp-4	4750			0		46
F25B4.2	2875			0		47
ins-32	454			0		47
dpy-14	1006			0		48
efn-2	2744			0		49
arrd-17	1000			0		49
Y55D5A.1	3419			0		54
gpa-14	3000			0		56
tax-4	1080			0		58
sre-1	1006			0		61
K10G6.4	2380			0		62
c33A12.4	604			0		62
tbb-4	3000			0		64
gpc-1	4200			0		64
jbts-14	3761			0		66
itf-81	1457			0		67
daf-28	3302			0		67

The *E*-value of a sequence is the expected number of sequences in a random database of the same size that would match the motifs as well as the sequence does.The position *P*-value of a match is the probability of a single random subsequence of the length of the motif scoring at least as well as the observed match. For each sequence, all motif occurrences with a position *P*-value <0.0001 are shown.

Sp family TFs in *C. elegans*, only SPTF-1 appears to be involved in the regulation of gene expression in ASJ neurons.

SPTF-1 displays characteristics of an ASJ terminal selector gene

If SPTF-1 is necessary for ASJ-specific gene expression, as a putative ASJ terminal selector gene, SPTF-1 should be expressed in these neurons, it should be autoregulated, and it should not interfere with the mitotic development of ASJ neurons because terminal selector genes code for postmitotically expressed TFs (Hobert 2010). Hence, we first aimed to determine the *sptf-1* expression pattern in transgenic animals harboring an *sptf-1* GFP reporter in which GFP is fused in-frame at the end of the second exon of *sptf-1*. Interestingly, among all the head neurons, SPTF-1 is exclusively expressed in ASJ neurons as shown both by colocalization with DsRed expressed in ASJ under the control of the *trx-1* promoter (Figure 4A) and by fluorescent DiI staining and DIC microscopy (Figure 4, B and E). Consistent with its expression in ASJ neurons, we have identified a putative ASJ motif within the *sptf-1* promoter region where both individual elements of the bipartite ASJ motif are relatively well conserved among *Caenorhabditis* species, while the 3 bp linker between them is more divergent (Figure 4F). In addition, SPTF-1 is also expressed in non-neuronal cells of the animal such as pharyngeal cells, rectal and intestinal cells, seam cells, and vulva epithelial cells (Figure 4C).

To analyze autoregulation of *sptf-1* expression in ASJ, we knocked down *sptf-1* using an RNAi construct directed at the last two *sptf-1* exons, thus targeting only the endogenous *sptf-1* gene but not the *sptf-1* GFP reporter. This approach abolished GFP expression in ASJ (Figure 4D), although the penetrance of this RNAi was variable and some animals displayed weak fluorescence in one or both ASJ neurons.

Finally, L1 larvae fed with *sptf-1* RNAi did not show DiI staining of the ASJ neurons although ASJ was present, as visualized by DIC microscopy (Figure 4E). This result demonstrates that *sptf-1* is not required for ASJ neurogenesis but instead regulates the expression of genes required for DiI dye uptake into these neurons.

Collectively, these data indicate that the *Sp* transcription factor family member SPTF-1 is an ASJ neuron terminal selector gene, probably acting by binding to the CG-rich element of the bipartite ASJ motif in *C. elegans*. To the best of our knowledge, this is the first biological function assigned to *C. elegans* SPTF-1.

ASJ motif is found in the promoters of many ASJ neuron-expressed genes

Next, we used the consensus ASJ motif from *trx-1*, *ssu-1*, and *sptf-1* promoters (Figure 4F) to search for the respective ASJ motifs in the promoters of all genes reported to be expressed in ASJ neurons using the MEME Suite Software

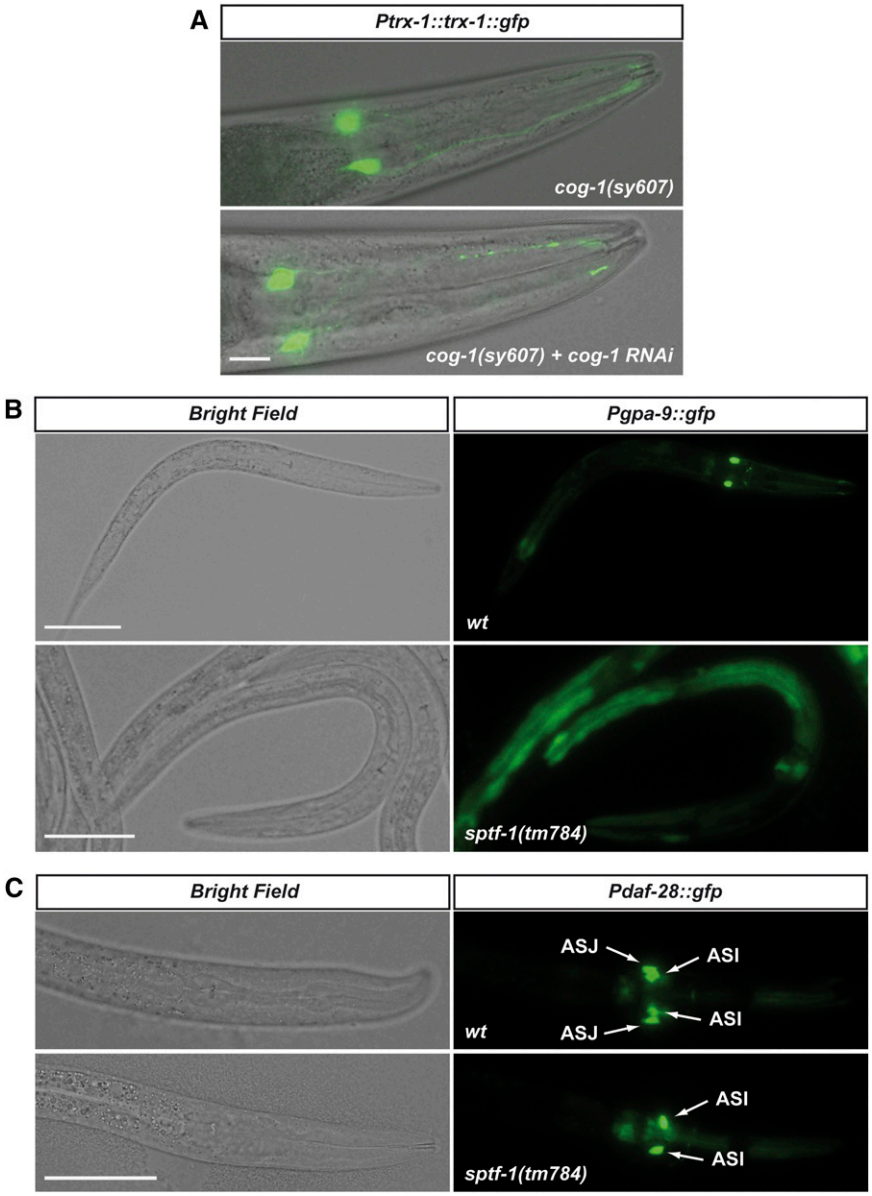

Figure 5 COG-1 is not required for ASJ specification whereas SPTF-1 regulates ASJ expression of DAF-28 and GPA-9. (A) *cog-1*(sy607) homozygous animals express *Ptrx-1*::*trx-1*::*gfp* reporter in ASJ. *cog-1* RNAi was used to exclude a possible effect of maternally contributed COG-1. Bar, 20 μm. (B and C) *sptf-1* is required for ASJ expression of *gpa-9* and *daf-28* transcriptional GFP reporters in *sptf-1(tm784)* L1 larvae. Note that *sptf-1* deficiency does not affect expression of *daf-28* in ASI neurons. Bar, 20 μm.

(http://meme.nbcr.net/meme/). As shown in Table 1, many of these ASJ-expressed genes have putative ASJ motifs within their promoters. Interestingly, among the genes with the highest score, we found *cog-1*, a homeobox transcription factor that has been implicated in *C. elegans* reproductive system development and ASE left/right asymmetry (Palmer *et al.* 2002; Chang *et al.* 2003). Given that the second element of the bipartite ASJ motif is an AT-rich sequence resembling a typical homeodomain-binding core (Berger *et al.* 2008), *cog-1* appeared as a good candidate to bind this second element. However, neither *cog-1* RNAi nor mutation suppressed ASJ labeling in a *Ptrx-1*::*trx-1*::*gfp* reporter strain, thus arguing against any regulatory role of this TF in ASJ-specific gene expression (Figure 5A).

To further experimentally validate *sptf-1* as a terminal selector gene for ASJ-expressed genes other than *trx-1*, *ssu-1*, and *sptf-1*, we analyzed ASJ labeling in *sptf-1* mutants expressing a *Pgpa-9*::*gfp* reporter expressed only in ASJ in the head (plus PVQ and PHB neurons in the tail) (Jansen *et al.* 1999) and a *Pdaf-28*::*gfp* reporter expressed in ASI and ASJ neurons (Li *et al.* 2003). In both cases, *sptf-1(tm784)* arrested L1 larvae were devoid of GFP expression in the ASJ neurons (Figure 5, B and C), further supporting *sptf-1* as a ASJ terminal selector gene. As previously described for *trx-1* and *ssu-1* GFP reporters, we also found that the *Pdaf-28*::*gfp* transgene, in some cases, is ectopically expressed in several head neurons other than ASI and ASJ (Figure S1D).

A homeobox transcription factor may bind the second element of the bipartite ASJ motif

The second element of the bipartite ASJ motif is an AT-rich sequence, including a typical homeodomain-binding core (Berger *et al.* 2008). As mentioned above, our RNAi screening, which included all predicted homeobox TFs, did not

result in any hit. This could either be due to the low efficiency of RNAi-mediated down-regulation, as neurons are usually quite refractory to feeding RNAi approaches (despite the use of an RNAi hypersensitive background, validated by the highly efficient *sptf-1* RNAi) or due to redundancies among different homeobox TFs. We also used the modENCODE tool (http://www.modencode.org) to identify TFs that potentially bind the *C. elegans trx-1* promoter. Based on this analysis, we selected homeobox TFs whose predicted binding sites resemble the ASJ motif. However, mutant alleles of the five candidates identified, *mab-5*, *zag-1*, *ceh-14*, *ceh-26*, and *ceh-30*, failed to abolish *Ptrx-1*::*gfp* expression in ASJ neurons (data not shown). Together, these data suggest that several homeobox TFs might redundantly bind to this second element of the bipartite ASJ motif, as has been reported for other homeobox and T-box TFs (Good *et al.* 2004; Furuya *et al.* 2005). Alternatively, it is possible that this AT-rich element is necessary for *sptf-1*-dependent ASJ expression but does not require binding of any transcription factor, as described for the GC-box-binding MIG1 zinc finger (Lundin *et al.* 1994).

Conclusions

In summary, we report here the identification of the *cis*-regulatory promoter motif responsible for gene expression in the ASJ neurons in *C. elegans*. This ASJ motif is bipartite and composed of CG-rich and AT-rich sequence elements separated by a divergent linker sequence. Given this sequence composition, we propose that one or more still-to-be-identified homeobox TF(s) might bind the AT-rich element of the bipartite ASJ motif. In turn, we have identified the *Sp* family member SPTF-1 as a transcription factor required for ASJ-specific gene expression, most likely by binding to the CG-rich element of the bipartite motif. Among all *C. elegans* neurons, SPTF-1 is expressed only in ASJ neurons where it functions as a terminal selector gene.

Acknowledgments

Some *C. elegans* strains were provided by the CGC, which is funded by the National Institutes of Health Office of Research Infrastructure Programs (P40 OD010440), and by the Japanese National Bioresource Project, which is funded by the Japanese Ministry of Education, Culture, Sport, Science and Technology. We thank Nuria Flames for advice and support and María Jesús Rodríguez-Palero and Francisco José Naranjo-Galindo for excellent technical assistance. This work was financed by grants to A.M.-V. from the Junta de Andalucía (Projects P07-CVI-02697 and P08-CVI-03629). Work in the laboratory of P.S. was supported by grants from the Swedish Research Council and the Torsten Söderberg Foundation. E.K. was supported by a grant from the European Union FP6 Marie Curie Research Training Network "EUrythron" MRTN-CT-2004-005499.

Literature Cited

Alcedo, J., and C. Kenyon, 2004 Regulation of C. elegans longevity by specific gustatory and olfactory neurons. Neuron 41: 45–55.

Bargmann, C. I., and H. R. Horvitz, 1991 Control of larval development by chemosensory neurons in Caenorhabditis elegans. Science 251: 1243–1246.

Bejerano, G., M. Pheasant, I. Makunin, S. Stephen, W. J. Kent *et al.*, 2004 Ultraconserved elements in the human genome. Science 304: 1321–1325.

Berger, M. F., G. Badis, A. R. Gehrke, S. Talukder, A. A. Philippakis *et al.*, 2008 Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell 133: 1266–1276.

Bertrand, V., and O. Hobert, 2009 Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in C. elegans. Dev. Cell 16: 563–575.

Bloomfield, K. L., S. A. Osborne, D. D. Kennedy, F. M. Clarke, and K. F. Tonissen, 2003 Thioredoxin-mediated redox control of the transcription factor Sp1 and regulation of the thioredoxin gene promoter. Gene 319: 107–116.

Boulin, T., J. F. Etchberger, and O. Hobert, 2006 Reporter gene fusions (April 5, 2006), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.106.1, http://www.wormbook.org.

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

Bulger, M., and M. Groudine, 2011 Functional and mechanistic diversity of distal transcription enhancers. Cell 144: 327–339.

Carroll, B. T., G. R. Dubyak, M. M. Sedensky, and P. G. Morgan, 2006 Sulfated signal from ASJ sensory neurons modulates stomatin-dependent coordination in Caenorhabditis elegans. J. Biol. Chem. 281: 35989–35996.

Chang, S., R. J. Johnston, Jr., and O. Hobert, 2003 A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans. Genes Dev. 17: 2123–2137.

Cornils, A., M. Gloeck, Z. Chen, Y. Zhang, and J. Alcedo, 2011 Specific insulin-like peptides encode sensory information to regulate distinct developmental processes. Development 138: 1183–1193.

Davidson, E. H., 2001 *Genomic Regulatory Systems*. Academic Press, The Netherlands.

Davidson, E. H., 2006 *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution*. Academic Press, The Netherlands.

de Bono, M., 2003 Molecular approaches to aggregation behavior and social attachment. J. Neurobiol. 54: 78–92.

Doitsidou, M., N. Flames, I. Topalidou, N. Abe, T. Felton *et al.*, 2013 A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in *C. elegans.* Genes Dev. 27: 1391–1405.

Etchberger, J. F., and O. Hobert, 2008 Vector-free DNA constructs improve transgene expression in C. elegans. Nat. Methods 5: 3.

Etchberger, J. F., A. Lorch, M. C. Sleumer, R. Zapf, S. J. Jones *et al.*, 2007 The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. Genes Dev. 21: 1653–1674.

Etchberger, J. F., E. B. Flowers, R. J. Poole, E. Bashllari, and O. Hobert, 2009 Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in C. elegans. Development 136: 147–160.

Fierro-Gonzalez, J. C., A. Cornils, J. Alcedo, A. Miranda-Vizuete, and P. Swoboda, 2011a The thioredoxin TRX-1 modulates the function of the insulin-like neuropeptide DAF-28 during dauer formation in Caenorhabditis elegans. PLoS ONE 6: e16561.

Fierro-Gonzalez, J. C., M. Gonzalez-Barrios, A. Miranda-Vizuete, and P. Swoboda, 2011b The thioredoxin TRX-1 regulates adult lifespan extension induced by dietary restriction in Caenorhabditis elegans. Biochem. Biophys. Res. Commun. 406: 478–482.

Flames, N., and O. Hobert, 2009 Gene regulatory logic of dopamine neuron differentiation. Nature 458: 885–889.

Furuya, M., H. Qadota, A. D. Chisholm, and A. Sugimoto, 2005 The C. elegans eyes absent ortholog EYA-1 is required

for tissue differentiation and plays partially redundant roles with PAX-6. Dev. Biol. 286: 452–463.
Good, K., R. Ciosk, J. Nance, A. Neves, R. J. Hill *et al.*, 2004 The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C. elegans embryos. Development 131: 1967–1978.
Hempel, N., H. Wang, E. L. LeCluyse, M. E. McManus, and M. Negishi, 2004 The human sulfotransferase SULT1A1 gene is regulated in a synergistic manner by Sp1 and GA binding protein. Mol. Pharmacol. 66: 1690–1701.
Hobert, O., 2008 Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. Proc. Natl. Acad. Sci. USA 105: 20067–20071.
Hobert, O., 2010 Neurogenesis in the nematode Caenorhabditis elegans (October 4, 2010), WormBook, ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.12.2, http://www.wormbook.org.
Jansen, G., K. L. Thijssen, P. Werner, M. van der Horst, E. Hazendonk *et al.*, 1999 The complete family of genes encoding G proteins of Caenorhabditis elegans. Nat. Genet. 21: 414–419.
Junion, G., M. Spivakov, C. Girardot, M. Braun, E. H. Gustafson *et al.*, 2012 A transcription factor collective defines cardiac cell fate and reflects lineage history. Cell 148: 473–486.
Kaczynski, J., T. Cook, and R. Urrutia, 2003 Sp1- and Kruppel-like transcription factors. Genome Biol. 4: 206.
Kim, K., R. Kim, and P. Sengupta, 2010 The HMX/NKX homeodomain protein MLS-2 specifies the identity of the AWC sensory neuron type via regulation of the ceh-36 Otx gene in C. elegans. Development 137: 963–974.
Lai, C. H., C. Y. Chou, L. Y. Ch'ang, C. S. Liu, and W. Lin, 2000 Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics. Genome Res. 10: 703–713.
Lee, Y. C., Y. Higashi, C. Luu, C. Shimizu, and C. A. Strott, 2005 Sp1 elements in SULT2B1b promoter and 5′-untranslated region of mRNA: Sp1/Sp2 induction and augmentation by histone deacetylase inhibition. FEBS Lett. 579: 3639–3645.
Levine, M., 2010 Transcriptional enhancers in animal development and evolution. Curr. Biol. 20: R754–R763.
Li, W., S. G. Kennedy, and G. Ruvkun, 2003 daf-28 encodes a C. elegans insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. Genes Dev. 17: 844–858.
Lundin, M., J. O. Nehlin, and H. Ronne, 1994 Importance of a flanking AT-rich region in target site recognition by the GC box-binding zinc finger protein MIG1. Mol. Cell. Biol. 14: 1979–1985.
Mello, C. C., J. M. Kramer, D. Stinchcomb, and V. Ambros, 1991 Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 10: 3959–3970.
Miranda-Vizuete, A., J. C. Fierro Gonzalez, G. Gahmon, J. Burghoorn, P. Navas *et al.*, 2006 Lifespan decrease in a Caenorhabditis elegans mutant lacking TRX-1, a thioredoxin expressed in ASJ sensory neurons. FEBS Lett. 580: 484–490.
Morgan, T. H., 1934 *Embryology and Genetics*. Columbia University Press, New York.
Nokes, E. B., A. M. Van Der Linden, C. Winslow, S. Mukhopadhyay, K. Ma *et al.*, 2009 Cis-regulatory mechanisms of gene expression in an olfactory neuron type in Caenorhabditis elegans. Dev. Dyn. 238: 3080–3092.
Ohta, A., T. Ujisawa, S. Sonoda, and A. Kuhara, 2014 Light and pheromone-sensing neurons regulates cold habituation through insulin signalling in Caenorhabditis elegans. Nat. Commun. 5: 4412.
Palmer, R. E., T. Inoue, D. R. Sherwood, L. I. Jiang, and P. W. Sternberg, 2002 Caenorhabditis elegans cog-1 locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development. Dev. Biol. 252: 202–213.
Perkins, L. A., E. M. Hedgecock, J. N. Thomson, and J. G. Culotti, 1986 Mutant sensory cilia in the nematode Caenorhabditis elegans. Dev. Biol. 117: 456–487.
Rankin, C. H., 2002 From gene to identified neuron to behaviour in Caenorhabditis elegans. Nat. Rev. Genet. 3: 622–630.
Reece-Hoyes, J. S., B. Deplancke, J. Shingles, C. A. Grove, I. A. Hope *et al.*, 2005 A compendium of Caenorhabditis elegans regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol. 6: R110.
Riddle, D. L., M. M. Swanson, and P. S. Albert, 1981 Interacting genes in nematode dauer larva formation. Nature 290: 668–671.
Sandelin, A., P. Bailey, S. Bruce, P. G. Engstrom, J. M. Klos *et al.*, 2004 Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. BMC Genomics 5: 99.
Schackwitz, W. S., T. Inoue, and J. H. Thomas, 1996 Chemosensory neurons function in parallel to mediate a pheromone response in C. elegans. Neuron 17: 719–728.
Shaye, D. D., and I. Greenwald, 2011 OrthoList: a compendium of C. elegans genes with human orthologs. PLoS ONE 6: e20085.
Suske, G., 1999 The Sp-family of transcription factors. Gene 238: 291–300.
Svingen, T., and K. F. Tonissen, 2006 Hox transcription factors and their elusive mammalian gene targets. Heredity (Edinb) 97: 88–96.
Swoboda, P., H. T. Adler, and J. H. Thomas, 2000 The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. Mol. Cell 5: 411–421.
Timmons, L., D. L. Court, and A. Fire, 2001 Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene 263: 103–112.
Ulm, E. A., S. F. Sleiman, and H. M. Chamberlin, 2011 Developmental functions for the Caenorhabditis elegans Sp protein SPTF-3. Mech. Dev. 128: 428–441.
Ward, A., J. Liu, Z. Feng, and X. Z. Xu, 2008 Light-sensitive neurons and channels mediate phototaxis in C. elegans. Nat. Neurosci. 11: 916–922.
Wenick, A. S., and O. Hobert, 2004 Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. Dev. Cell 6: 757–770.
White, J. G., E. Southgate, J. N. Thomson, and S. Brenner, 1986 The structure of the nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. Lond. B Biol. Sci. 314: 1–340.
Wimmer, E. A., G. Frommer, B. A. Purnell, and H. Jackle, 1996 buttonhead and D-Sp1: a novel Drosophila gene pair. Mech. Dev. 59: 53–62.
Winkelbauer, M. E., J. C. Schafer, C. J. Haycraft, P. Swoboda, and B. K. Yoder, 2005 The C. elegans homologs of nephrocystin-1 and nephrocystin-4 are cilia transition zone proteins involved in chemosensory perception. J. Cell Sci. 118: 5575–5587.
Woolfe, A., M. Goodson, D. K. Goode, P. Snell, G. K. McEwen *et al.*, 2005 Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol. 3: e7.
Wray, G. A., 2007 The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8: 206–216.
Zhang, F., A. Bhattacharya, J. C. Nelson, N. Abe, P. Gordon *et al.*, 2014 The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development 141: 422–435.

Communicating editor: P. Sengupta

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.176172/-/DC1

Cis- and *Trans*-Regulatory Mechanisms of Gene Expression in the ASJ Sensory Neuron of *Caenorhabditis elegans*

María González-Barrios, Juan Carlos Fierro-González, Eva Krpelanova, José Antonio Mora-Lorca, José Rafael Pedrajas, Xenia Peñate, Sebastián Chavez, Peter Swoboda, Gert Jansen, and Antonio Miranda-Vizuete

DOI: 10.1534/genetics.115.176172

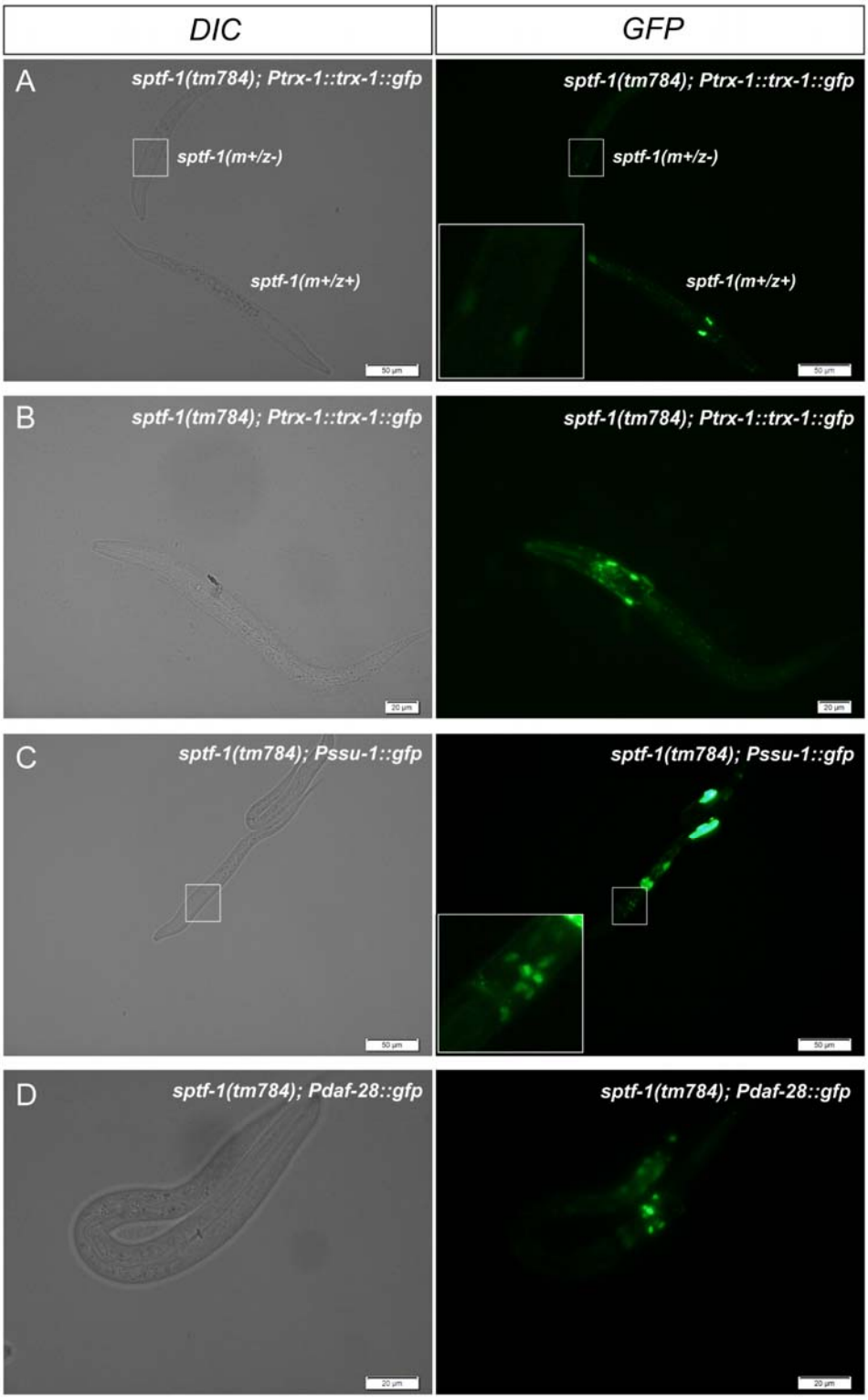

Figure S1 ***sptf-1*** **deficiency causes anomalous ectopic neuronal expression of fluorescence reporters of genes normally expressed in ASJ neurons. A)** In some cases, *sptf-1* arrested L1 larvae display very weak GFP labeling in ASJ neurons, possibly due to maternal effect. **B,C,D)** In other cases, transgenic animals expressing ASJ fluorescence reporters display signal in several additional neurons in the head area.

Table S1 is available for download at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.176172/-/DC1