Published online 8 September 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 1* ***53–64***
doi:10.1093/nar/gkr690

Fine tuning of RFX/DAF-19-regulated target gene expression through binding to multiple sites in *Caenorhabditis elegans*

Jeffery S. C. Chu[1], Maja Tarailo-Graovac[1], Di Zhang[1], Jun Wang[1], Bora Uyar[1,2], Domena Tu[1], Joanne Trinh[1], David L. Baillie[1] and Nansheng Chen[1,2,*]

[1]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby and [2]CIHR/MSFHR Bioinformatics Training Program, Vancouver, BC, Canada

Received June 7, 2011; Revised July 26, 2011; Accepted August 6, 2011

ABSTRACT

In humans, mutations of a growing list of regulatory factor X (RFX) target genes have been associated with devastating genetics disease conditions including ciliopathies. However, mechanisms underlying RFX transcription factors (TFs)-mediated gene expression regulation, especially differential gene expression regulation, are largely unknown. In this study, we explore the functional significance of the co-existence of multiple X-box motifs in regulating differential gene expression in *Caenorhabditis elegans*. We hypothesize that the effect of multiple X-box motifs is not a simple summation of binding effect to individual X-box motifs located within a same gene. To test this hypothesis, we identified eight *C. elegans* genes that contain two or more X-box motifs using comparative genomics. We examined one of these genes, F25B4.2, which contains two 15-bp X-box motifs. F25B4.2 expression in ciliated neurons is driven by the proximal motif and its expression is repressed by the distal motif. Our data suggest that two X-box motifs cooperate together to regulate the expression of F25B4.2 in location and intensity. We propose that multiple X-box motifs might be required to tune specific expression level. Our identification of genes with multiple X-box motifs will also improve our understanding of RFX/DAF-19-mediated regulation in *C. elegans* and in other organisms including humans.

INTRODUCTION

Regulatory factor X (RFX) is an evolutionarily conserved DNA binding protein family that has been identified in organisms ranging from single cellular eukaryotes, including the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe*, to humans (1). All RFX transcription factors (TFs) contain a single DNA binding domain (DBD) that is very well conserved, showing over 40% identity between yeast, nematodes and mammals and close to 100% identity in amino acid positions that are in direct contact with RFX DBDs (2,3). We have found that different organisms in the tree of life can have very different number of RFX genes. Non-metazoans except the choanoflagellates (e.g. *Monosiga brevicollis*) have either one RFX gene or none (2). In contrast, all metazoans have multiple RFX genes, except the nematodes which only have a single RFX gene. For example, *Caenorhabditis elegans* genome has a single RFX gene called *daf-19*, which is expressed in ciliated neurons (4). It has been proposed that nematodes could have lost some RFX genes in evolution (1,2). The first RFX TF in *Drosophila melanogaster*, dRFX (CG6312), was identified through a homology search for the RFX DBD and the second one, dRFX2, was identified through yeast-one-hybrid (Y1H) screening for transcription factors that bind to a putative promoter sequence (5,6). However, this gene sequence does not match any genomic region in the sequenced *D. melanogaster* genome. Because the *D. melanogaster* heterochromatin regions have not been fully sequenced due to heavy repetitive sequences (7), it cannot be ruled out that dRFX2 resides in these unsequenced regions. Because dRFX2 phylogenetically clusters with fungal RFX genes, it has been proposed that dRFX2 is in fact a fungal gene (2). Nevertheless, perfect match in any sequenced fungal species was not found (2). In a recent comparative genomics study, we have annotated the gene CG9727 in *D. melanogaster* as the third RFX TF dRFX1, whose DBD shows high similarity to that of human RFX5 (2). Seven RFX TFs have been uncovered in humans as well as all other mammals, including RFX6 and RFX7 that were recently identified in

*To whom correspondence should be addressed. Tel: +1 778 782 4823; Fax: +1 778 782 5583; Email: chenn@sfu.ca

our laboratory (8). RFX3 in mammals has been found to be crucial for proper primary cilia development in embryo nodal cells (9), brain ependymal cells (10) and pancreatic endocrine cells (11). Recently, RFX4 was also shown to be important for cilia formation in the neural tube (12). Mutations in many of their target genes have been associated with an expanding array of devastating human disease conditions, including polycystic kidney disease (13,14) and Bardet-Biedl syndrome (BBS) (15). RFX5 regulates major histocompatibility complex class II (MHC II) gene expression in the immune system (16). RFX6, which we recently identified in the human genome, is almost exclusively expressed in the human pancreatic islets (8) and plays a role in the insulin production (17,18).

Accumulating evidence suggests that RFX genes regulate the transcription of ciliary genes in metazoans. The first critical evidence linking RFX TFs and the ciliary genes was reported by Swoboda *et al*. (4). The authors cloned *daf-19* in the nematode *C. elegans* and found that it is the first and only RFX gene in *C. elegans*. They showed that in the absence of functional DAF-19, ciliated neurons in *C. elegans* lost their cilia and displayed chemosensory defects (Che), dye filling defect (Dyf) and constitutive dauer formation (Daf-c) (4). Furthermore, they demonstrated that DAF-19 regulates the expression of ciliary genes, including *che-2*, *osm-1*, *osm-6* and many BBS genes through binding to a DNA element called the X-box motif, which was first identified as binding site for human RFX5 (4,19). Later, it was discovered that many ciliary genes in the fruit fly *D. melanogaster* are also regulated by RFX genes, including *CG15161*, the homolog of *dyf-6* (20).

X-box motif, the binding motif of RFX DBDs, has been found to be highly conserved as well. Many validated instances of X-box motifs in yeast, *C. elegans* and humans are 14-bp in size. Because of their large size, X-box motifs have been used as a ciliary gene indicator in genomics and bioinformatics projects. Efimenko *et al*. (21) searched the *C. elegans* promoter regions (defined here as 1000-bp genomic region upstream of the start codon) for candidate X-box motifs that resemble an 'average consensus X-box motif' and identified 730 potential DAF-19 target genes in *C. elegans*. Independently, Blacque *et al*. (22) identified 53 putative DAF-19 target genes in *C. elegans* through searching for the presence of putative X-box motifs in promoter regions (defined as 1500-bp genomic region upstream of the start codon) and comparing relative gene expression in four different tissues. Taking advantage of the availability of two newly sequenced genomic sequences in *Caenorhabditis* genus, *C. briggsae* and *C. remanei*, our laboratory searched for X-box motifs in the promoter regions (defined as 2000-bp genomic region upstream of the start codon) of orthologous genes in all three species and predicted 93 candidate DAF-19 regulated genes (23), including *dyf-5*. The putative X-box motifs identified in these three studies all show resemblance to known X-box motifs.

Many RFX target genes have been found to contain more than one X-box motif. For example, RFX TF in *S. cerevisiae* negatively regulates many ribonucleotide reductase genes (e.g. RNR2, RNR3 and RNR4) through a combination of strong X-box motifs and weak X-box motifs (24). In humans, RFX1 represses MAP1A in non-neuronal cells by binding to two X-box motifs in the first exon (25). However, how multiple X-box motifs function in the same gene has never been examined in animals *in vivo*.

The goal of this project is to identify ciliary genes in *C. elegans* that may be regulated by DAF-19 through multiple X-box motifs. *C. elegans* was chosen as a model system for this study because of its compact genome. More importantly, *C. elegans* has been effectively used to identify and characterize ciliary genes. Many ciliary genes in *C. elegans* have readily identifiable orthologs in humans (2,4,19), which makes this study useful for understanding the function and regulatory mechanism of RFX TFs in humans.

MATERIALS AND METHODS

Strains used

Worm strains are maintained using standard procedures in 20°C unless otherwise noted (26). The following strains were used in this study: DR86 *daf-19(m86)*, EG5003 *unc-119(ed3) III; cxTi10882 IV*, JT204 *daf-12(sa204)*, JT6924 *daf-12(sa204); daf-19(m86)*. Strains generated in this study are listed below.

Strain name	Allele	Description
JNC20	*unc-119*(ed3); dotSi1 [prF25B4.2::mCherry, *Cb-unc-119*(+)] IV	F25B4.2 Wild-type promoter
JNC21	*unc-119*(ed3); dotSi2 [prF25B4.2::mCherry del(−140), *Cb-unc-119*(+)] IV	Proximal deletion
JNC22	*unc-119*(ed3); dotSi3 [prF25B4.2::mCherry del(−190), *Cb-unc-119*(+)] IV	Distal deletion
JNC29	*unc-119*(ed3); dotSi10 [prF25B4.2::mCherry del(-140) del(190), *Cb-unc-119*(+)] IV	Double deletion
JNC33	*daf-19*(m86); dotSi2	Proximal deletion in *daf-19* background
JNC34	*daf-19*(m86); dotSi10	Double deletion in *daf-19* background
JNC36	*daf-19*(m86); dotSi1	Wild-type promoter in *daf-19* background
JNC37	*daf-19*(m86); dotSi3	Distal deletion in *daf-19* background
JNC23	*unc-119*(ed3); dotSi4 [prM04C9.5::mCherry, *Cb-unc-119*(+)] IV	*dyf-5* wild-type promoter
JNC31	*unc-119*(ed3); dotSi4 [prM04C9.5::mCherry replace −285 to −271 →gtcctcacaagtaac, *Cb-unc-119*(+)] IV	*dyf-5* promoter replaced with distal motif
JNC35	*unc-119*(ed3); dotSi4 [prM04C9.5::mCherry replace −285 to −271 →gtctccaatggcaac, *Cb-unc-119*(+)] IV	*dyf-5* promoter replaced with proximal motif

Genomic data and gene model improvement

Genomics DNA data for all four *Caenorhabditis* species were obtained from WS204 version of WormBase

(ftp.wormbase.org/). The gene set for *C. elegans* was obtained from release WS204. The gene set for *C. briggsae*, *C. remanei* and *C. brenneri* were generated by genBlastG (27). *Caenorhabditis elegans* proteins (20 173) were used as input for genBlastG. These sequences represent the longest transcript if more than one alternative transcript exits. genBlastG returns 264 411 gene models for *C. briggsae*, 319 750 for *C. remanei* and 425 947 for *C. brenneri*. Many of these gene models are redundant gene models due to multiple orthologs (such as gene families or tandem gene duplications) in *C. elegans*. A filtering procedure was used so that each genomic region would contain only one gene model with the highest sequence percent identity (PID) to the query. The filtering procedure was carried out as follows: (i) all predictions are sorted by PID in decreasing order. (ii) For each two overlapping model, if the overlapping region is >5% of the length for either gene, then the model with higher PID is kept and the model with lower PID is filtered out. (iii) To ensure the quality of the gene set, we only kept gene models that show PID ≥40% to the query. The filtering procedure resulted in 16 577 gene models for *C. briggsae*, 18 426 for *C. remanei* and 23 473 for *C. brenneri*. In the last step, we combined these gene models with the current WormBase models to generate a hybrid gene set. In the hybrid set, genBlastG's predicted models replace corresponding WormBase gene models if genBlastG's prediction shows at least 2% improvement in PID. The final gene models were uploaded in GFF3 format to a MYSQL server and visualized on Generic Genome Browser (28). Mapping of orthologous relationships between genes in *C. elegans* and the other three species were generated by Inparanoid (29).

X-box motif search using TFMscan

We generated a position weighted matrix (PWM) for the left 6-bp and the right 6-bp based on 31 validated X-box motifs (Table 1). The PWM for each half was used by TFMscan to probe the promoter regions. A putative X-box motif was predicted by combining a left half and a right half while allowing 2–3 nt in between. TFMscan was executed with $P = e^{-5}$ (30).

Construct generation and cloning

Deletion constructs were made by standard site-directed mutagenesis and PCR stitching methods (31). Briefly, primers were designed to contain the particular deletion. See below for a list of primers used and their sequences. A left fragment was amplified using Primer A and Primer DeletionR (either distal or proximal). A right fragment was amplified using Primer DeletionF (either distal or proximal) and Primer B. The left and right fragment is stitched together using Primer A* and B. The mCherry was amplified using Primer C and D from pCFJ190 (A gift from E.M. Jorgensen). The final stitching between the promoter fragment and mCherry was done using Primer A* and D*. Primer A* contains SbfI site and Primer D* contains SpeI site. The construct and the plasmid pCFJ178 were cut using the respective restriction enzymes in 37°C for 2.5 h. The construct was ligated into the linearized plasmid overnight at room temperature. The final ligation reaction was transformed into DH5α cells by electroporation. The transformants were plated onto lysogeny broth (LB)-Ampicillin plates. Living colonies were picked to grow in a 5 ml of LB broth and the DNA was extracted using Qiagen Mini-prep kit (Cat#:27104).

All the primers used in this study are listed below.

Primer name	Primer sequence
F25B4.2_A	CAAAATTACCTATCGCACTACGTT
F25B4.2_A*	CCTGCAGGCCTGCAGGAAGCTGA AACGTCGGAGATAATAC
F25B4.2_B	TATCTTCTTCACCCTTTGAGACCA TCATCCACGATTAATCTGAAA CTCA
M04C9.5_A	CCTGCAGGCCTGCAGGAATTGAA TTAGCCGCGGAGC
M04C9.5_B	TATCTTCTTCACCCTTTGAGACCA TGGCTTCTTGCCCTTATATTT TCC
mCherry_C	ATGGTCTCAAAGGGTGAAGA
mCherry_D	GGCCTCTTCGCTATTACGC
mCherry_D*	ACGACGGCCAGTGAATTATCACT AGTACTAGT
F25B4.2_deletionF_distal	CACTTTTCAATTCGAAATGTCATG GGCGTTG
F25B4.2_deletionR_distal	CCATGACATTTCGAATTGAAAAG TGTCGAAATTCTTAGAG
F25B4.2_deletionF_proximal	GGCGCCACTGAAACCCGCATTTT AAACTCCAT
F25B4.2_deletionR_proximal	CGGGTTTCAGTGGCGCCGTGGCG ACA
M04C9.5_replaceF_distal	GTCCTCACAAGTAACTGTCTGTTA CACCCTTTTCTC
M04C9.5_replaceR_distal	GTTACTTGTGAGGACCAAGAGCA AACGGCGGAG
M04C9.5_replaceF_proximal	GTCTCCAATGGCAACTGTCTGTTA CACCCTTTTCTC
M04C9.5_replaceR_proximal	GTTGCCATTGGAGACCAAGAGCA AACGGCGGAG
ChrIV-R	TGTTTACTAGACCGGGGCTC
mCherry-genoF	AAAACCGCACACAAAATACC
178-genoF	TCCCCATTTCACCAGAGAAC

MosSCI

DNA purified from the transformation was used directly for injection. The injection mix for MosSCI was made as suggested from the literature: pJL43.1 (50 ng/μl), purified plasmid (50 ng/μl), pGH8 (10 ng/μl), pCFJ90 (2.5 ng/μl), pCFJ104 (5 ng/μl). The mix was injected into EG5003 worms. Worms that move and show none of the mCherry markers were individually plated. To confirm an insertion, we performed PCR with primers ChrIV-F, mCherry-genoF and 178-genoR to genotype individual mothers. A worm with homozygous insertion would have a single band at ~2.2 kb; a worm with no insertion would have a single band at ~4 kb; a worm with heterozygous insertion would have both bands.

Dye filling assay

The methods for dye filling was adapted from Worm Atlas (32). Briefly, we washed one plate of mixed population using 1 ml of M9 buffer. Then, we collected worms using

centrifugation at 1500 rpm for 1 min and removed the supernatant. Then, we resuspended the worms in 1 ml of M9 buffer mixed with 5 µl of 2 mg/ml DiO (Molecular Probes, Cat#:D275). To allow the worms to take up the dye, we covered the tubes in tin foil and slowly shook at room temperature for 2 h. The worms were spun down again and transferred to a fresh seeded plate to allow the dye to pass through the gut. The worms were washed and spun as before just prior to transferring worms to the glass slide.

Genetic crosses

We obtained males for each strain containing the Mos insert by heat shocking 30 L4 hermaphrodites in 33°C for 4 h. We crossed four males with Mos insert to two *daf-19* (DR86) L4 hermaphrodites. In all, 15 hermaphrodite F1s were selected randomly and individually plated. The genotype of the *daf-19* gene in these F1s was confirmed by Tetra-ARM PCR (33). To find homozygous Mos insertion and homozygous *daf-19* mutation, we individually plated 200 F2s and screened for dauer phenotype as 85% of *daf-19* worms enter the dauer stage even in favorable condition (34). Candidates were screened and confirmed by genotyping.

Microscopy visualization

Worms were immobilized using azide on 3% agarose pad. Images were captured under Zeiss spinning disc confocal microscope (Zeiss Axio Observer.Z1) equipped with Hamamatsu ImagEM camera. Image capture and visualization were performed using Volocity software (www.improvision.com).

RESULTS

Search for putative *C. elegans* genes containing multiple X-box motifs by comparative genomics

Our search strategy involves finding genes that have conserved X-box motifs within 500-bp of promoter region in four *Caenorhabditis* species. We chose the 500-bp search space because all validated X-box motif identified to date reside in this region (Table 1). In order to obtain the 500-bp promoter region, we first performed gene annotation on *C. briggsae*, *C. remanei* and *C. brenneri* using genBlastG (see 'Materials and Methods' section). The *C. briggsae* gene models have not been examined since its publication (35) while the genomes of *C. remanei* and *C. brenneri* have not been published.

With the revised gene set generated from genBlastG, we employed TFMscan (30) to find X-box motifs within 500-bp upstream of the start codon. We focused on searching 14-bp and 15-bp X-box motifs since all validated X-box motifs identified to date are either 14-bp or 15-bp (see 'Materials and Methods' section, Table 1). We tested the sensitivity of TFMscan and it was able to capture 30 out of 31 validated X-box motifs

Table 1. Validated X-box motifs

Gene name	Sequence name	Distance from ATG	HMMER score	X-box sequence	Reference
che-13	F59C6.7	−74	5.33	GTTGCTATAGCAAC	(14,21)
xbx-1	F02D8.3	−79	7.73	GTTTCCATGGTAAC	(4,21,40)
xbx-2	D1009.5	−77	7.98	GTTGCCATGACAAC	(21,22)
xbx-3	M04D8.6	−97	5.72	GTTGTCTTGGCAAC	(21)
xbx-4	C23H5.3	−82	7.98	GTTGCCATGACAAC	(21)
xbx-5	T24A11.2	−121	6.84	GTCTCCATGACAAC	(21)
xbx-6	F40F9.1	−151	7.53	GTTTCCATGGAAAC	(21)
xbx-7	R148.1	−69	4.56	GTCACCATAGGAAC	(21)
	ZK328.7	−89	6.51	GTTACCATGGCAAT	(22)
bbs-9	C48B6.8	−81	7.53	GTTTCCATGACAAC	(22)
che-11	C27A7.4	−85	7.04	ATCTCCATGGCAAC	(21)
odr-4	Y102E9.1	−200	4.09	ATCGTCATGGTAAC	(21)
osm-5	Y41G9A.1	−115	6.92	GTTACTATGGCAAC	(21,36,37)
nhr-44	T19A5.4	−76	6.91	GTCTTCATGGCAAC	(21)
nph-1	M28.7	−77	5.57	GTTGCCAGGGGCAAC	(41)
nph-4	R13H4.1	−168	5.93	ATTTCCATGACAAC	(41)
nud-1	F53A2.4	−263	3.81	GTATCCATGGGAAC	(21)
dyf-2	ZK520.3	−140	5.84	GTTACCAAGGCAAC	(42)
osm-6	R31.3	−100	6.13	GTTACCATAGTAAC	(4,21,43)
dyf-3	C04C3.5	−88	4.22	GTTTCTATGGGAAC	(44,45)
	Y110A7A.20	−60	4.26	GTCTCTATAGCAAC	(22)
che−2	F38G1.1	−117	7.05	GTTGTCATGGTGAC	(4,21,46)
osm-1	T27B1.1	−86	5.27	GCTACCATGGCAAC	(4,21,34,47)
bbs-1	Y105E8A.5	−99	5.45	GTTCCCATAGCAAC	(21,48,49)
bbs-2	F20D12.3	−94	6.31	GTATCCATGGCAAC	(21,48,49)
bbs-5	R01H10.6	−65	8.64	GTCTCCATGGCAAC	(21,50)
bbs-7	Y75B8A.12	−94	6.92	GTTGCCATAGTAAC	(21,48,49)
bbs-8	T25F10.5	−84	4.22	GTACCCATGGCAAC	(21,48,49)
tub-1	F10B5.4	−183	5.25	ATCTCCATGACAAC	(21,51)
che-12	B0024.8	−767		ATCAGCTTGAAAAC	(52)
dyf-5	M04C9.5	−285	5.95	GTTACCATAGAAAC	(23,53)

An X-box motif containing gene is considered validated when the expression of the gene is dependent on the X-box motif (mutagenesis study) or on DAF-19 (DAF-19 knock out study).

Table 2. *Caenorhabditis elegans* genes with multiple putative X-box motifs within 500-bp promoter region

Sequence name	Gene name	Description	Putative X-box sequences	X-box position	Expression in ciliated neuron
C30F2.2		Similar to ARD GTP-binding proteins	gtcatccaggaaaac	−27	SAGE
			atcaccaatagcaac	−80	
C53B4.7	*bre-1*	Homologous to GDP-mannose 4, 6-dehydratases	gtttccttggcaaa	−423	GFP, SAGE
			atttcccatagaaac	−464	
F17A9.4		Homologous to NADH:flavin oxidoreductase	gtttgcccaacaac	−98	SAGE
			gtttctcaggcaac	−233	
F25B4.2		Homologous to Pellino	gtctccaatggcaac	−149	GFP, SAGE
			gtcctcacaagtaac	−199	
F32B6.9		Homologous to Bestrophin	gtctccttgacaac	−134	GFP, SAGE
			gtttctttgataac	−241	
			gttatcaaagaaac	−273	
H01G02.2		Homologous to CDK-Activating Kinase	gtctccatgacaac	−160	GFP, SAGE
			gttggcacggtaac	−234	
			gttaccgtgccaac	−280	
T13C2.7		Uncharacterized	gtcatctcagataac	−80	SAGE
			gtcacctaggaaac	−144	
Y41G9A.1	*osm-5*	Part of the intraflagellar transport	atctccatgacaac	−183	GFP, SAGE
			gtcgtcttggagac	−270	

Positions are listed as distance from the translational start site.

in *C. elegans*, which shows our search strategy is sensitive to identify *bona fide* X-box motifs. We searched 500-bp promoter region of every gene in *C. elegans* and found 91 genes that contain two or more putative X-box motifs. To further narrow down our candidate list, we asked which of the 91 genes have conserved X-box motif in *C. briggsae*, *C. remanei* and *C. brenneri*. Our query returned eight genes with such characteristics (Table 2).

We explored whether these eight genes show expression in ciliated neurons based on existing expression data. We found five genes with anatomical expression pattern data using GFP (Table 2). One of the genes, *osm-5*, was already characterized as a ciliary gene (21,36,37). Existing GFP reporter construct expression showed that F25B4.2, H01G02.2, C53B4.7 and F32B6.9 have neuronal expression (38,39). SAGE data further demonstrated that these genes are expressed in many ciliated neurons, including ASER and AFD neurons. The remaining three genes, though not having anatomical expression pattern data, show ciliated neuron expression based on SAGE data.

F25B4.2 is a conserved gene that harbors X-box motifs in four *Caenorhabditis* species but not in *C. elegans*

The expression pattern of F25B4.2 shown by GFP and SAGE tags indicates its potential for DAF-19 regulation. F25B4.2 contains two putative 15-bp X-box motifs in the upstream region, one located at 199-bp and the other at 149-bp upstream of the start codon (Figure 1a, Table 2). For convenience, we named the X-box motif located at −199 the distal motif and the X-box motif located at −149 the proximal motif. However, F25B4.2 was not identified as a DAF-19 target gene in any of the previous three genome-wide searches (21–23), likely due to its 15-bp X-box motifs that are different from the 14-bp consensus used. Among the two, the proximal motif displays higher conservation especially in the last six nucleotides where it is identical to many known X-box motifs (Figure 1b). However, these two motifs differ from the consensus at the third nucleotide (consensus = T) and at the eighth nucleotide (consensus = T). A single 15-bp X-box motif was also identified in three other *Caenorhabditis* species (Figure 1b).

F25B4.2 is conserved in four sequenced *Caenorhabdits* species with >80% identity at the protein level (Figure 2). The presence of X-box in upstream region of a gene usually suggests regulation by DAF-19 (4). F25B4.2 protein sequence shows ~40% identity to human Pellino gene family. Pellino proteins are E3 ligases known to participate in balancing inflammatory response (56). Pellino proteins interact with IRAK and mediate NFkB nuclear translocation to promote activation of pro-inflammatory genes (57,58). Pellino1 is suggested to play a part in TGF-β pathways to promote anti-inflammatory response preventing hyperactivation of inflammatory response (59,60).

Functional analysis of proximal and distal motifs

To examine whether the proximal or the distal motif is functional, we replaced the endogenous X-box motif of a known DAF-19 target gene promoter with either motif. If the motifs are functional, then it should drive ciliated neuron expression. We chose the promoter of *dyf-5* for this experiment because *dyf-5* was identified previously to express exclusively in ciliated neurons in a DAF-19-dependent manner (23,53). *dyf-5* expression is also X-box motif dependent such that its expression is completely abolished when the 14-bp X-box motif is removed (Wang,J. and Chen,N., unpublished data). We replaced the endogenous X-box motif in *dyf-5* promoter region (2000-bp upstream) with either the proximal motif or the distal motif and fused the promoter with mCherry. A single copy of the construct is integrated into the *C. elegans* genome at the Mos site cxTi10882 in chromosome IV using MosSCI method (61).

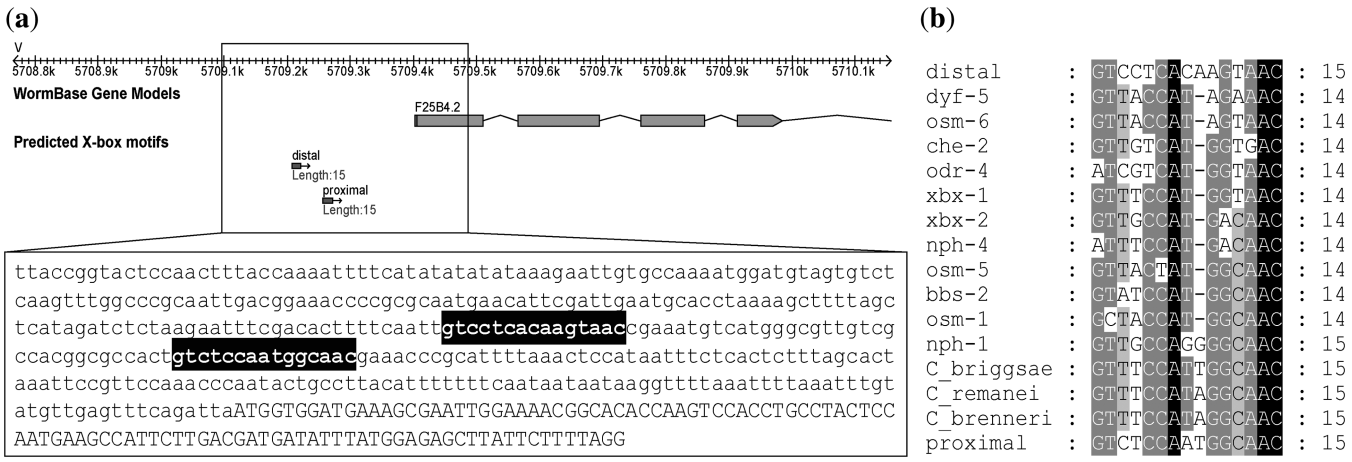

Figure 1. (**a**) The location and sequence of the putative X-box motifs upstream of F25B4.2. (**b**) The alignment of the putative proximal and distal motifs to known X-box motifs. Also include in the alignment are the putative X-box motifs in the orthologs of F25B4.2 in three other *Caenorhabditis* species labeled as *C. briggsae*, *C. remanei* and *C. brenneri*.

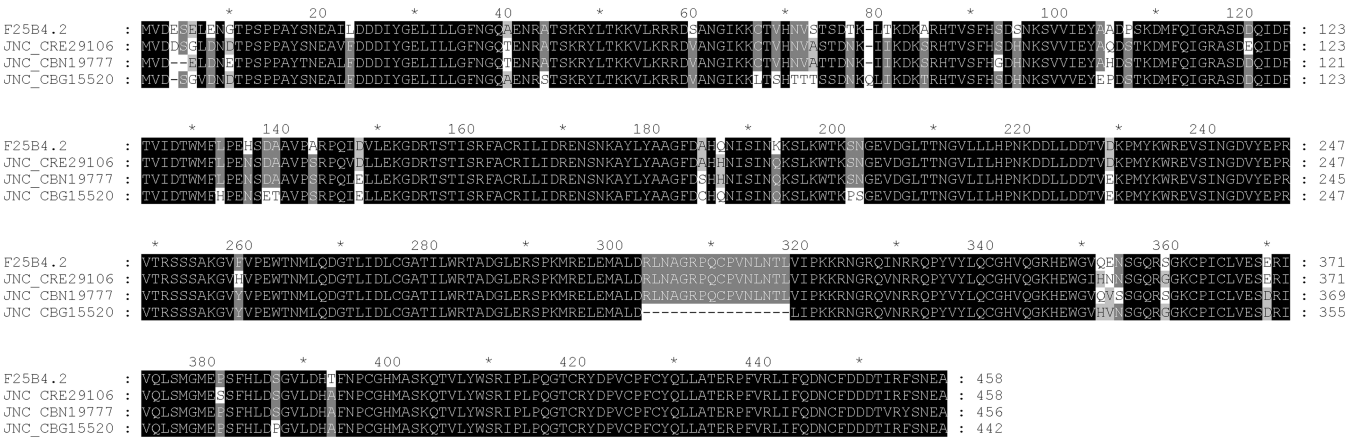

Figure 2. The alignment of F25B4.2 protein sequence with its orthologs in the other four *Caenorhabditis* species. CBG = *C. briggsae*, CRE = *C. remanei*, CBN = *C. brenneri*. The sequences were aligned using ClustalW (54) and visualized using GeneDoc (55).

This Mos element is located in an intergenic region with the flanking genes pointing toward each other. Hence, this location is not likely to have functional elements disrupted after reporter gene insertion. Insertion at the Mos site is confirmed by genotyping (Figure 3). The single copy-inserted construct using the wild-type *dyf-5* promoter is able to reproduce the ciliated neuron expression as reported before (Figure 4: *dyf-5* promoter) (23). In comparison to the endogenous X-box motif in *dyf-5*, the proximal element taken from F25B4.2 can drive robust expression in the amphid neurons and somewhat lower expression in the labial neurons (Figure 4: *dyf-5* promoter + proximal), suggesting that the proximal motif is a functional X-box motif. Interestingly, the distal element also drives mCherry expression in ciliated neurons in *C. elegans* but the expression level is greatly reduced (Figure 4: *dyf-5* promoter + distal). Our observations demonstrated that both motifs are functional motifs with the proximal motif able to drive much stronger expression than the distal motif.

F25B4.2 is expressed in ciliated neurons in a DAF-19-dependent manner

Since the proximal motif can be used to replace endogenous X-box motifs in the *dyf-5* promoter to drive ciliated neuron expression, we hypothesized that F25B4.2 expression depends on DAF-19. To test this hypothesis, we examined whether the promoter of F25B4.2 drives expression in ciliated neurons and whether its expression is dependent on DAF-19. We constructed a *C. elegans* strain carrying a single copy mCherry transgene driven by a 3000-bp promoter region upstream of F25B4.2 using MosSCI method (61). Observation of mCherry signals indicates that F25B4.2 is expressed in ciliated neurons (Figure 5). Dye-filling method with DiO in *C. elegans* allows six pairs of amphid neurons and two pairs of phasmid neurons to be filled with dye. Detailed analysis of F25B4.2 expression using dye-filling shows that F25B4.2 drives gene expression in ciliated neurons, including ASK, ADL, ASI, ASH, ASJ, PHA and PHB

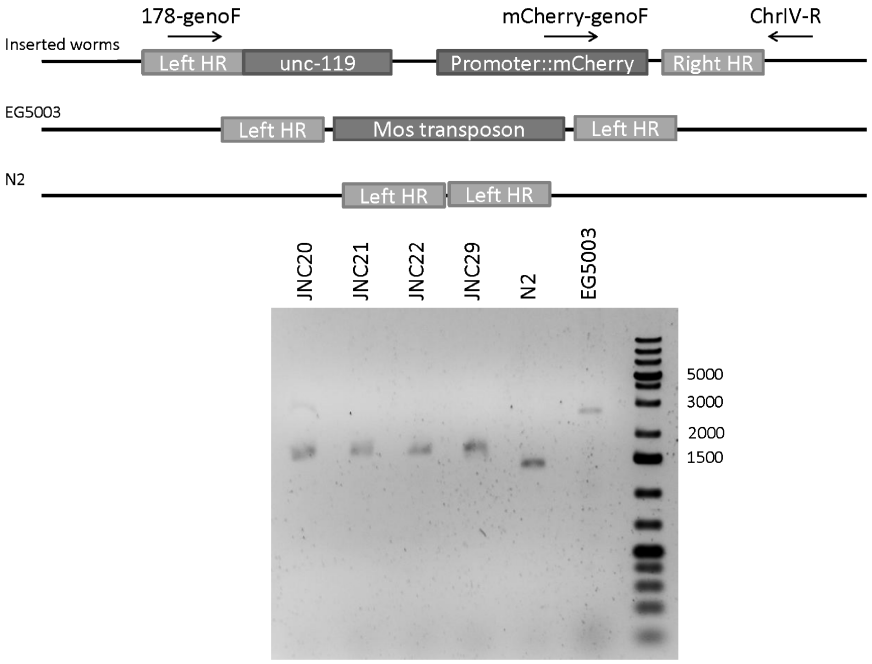

Figure 3. Genotyping of stably integrated strains. The insertion site on chromosome IV is depicted by the diagram on the top while an agarose gel showing the genotyping results on the bottom. The primers used for genotyping are also indicated by the arrows. Primer mCherry-genoF can only hybridize to inserted worms and not EG5003 and N2. The expected band sizes for inserted worms are 8312 bp from 178-genoF→ChrIV-R and 1564 bp from mCherry-genoF→ChrIV-R. The expected band size for EG5003 is 2700 bp [Mos1 is ~1280 bp (62–64)]. The expected band size for N2 is 1420 bp from 178-genoF→ChrIV-R. The gel image shows homozygous insertion for JNC20, 21, 22 and 29 as well as EG5003 and N2 as controls. The number on the right-hand side indicates the ladder positions.

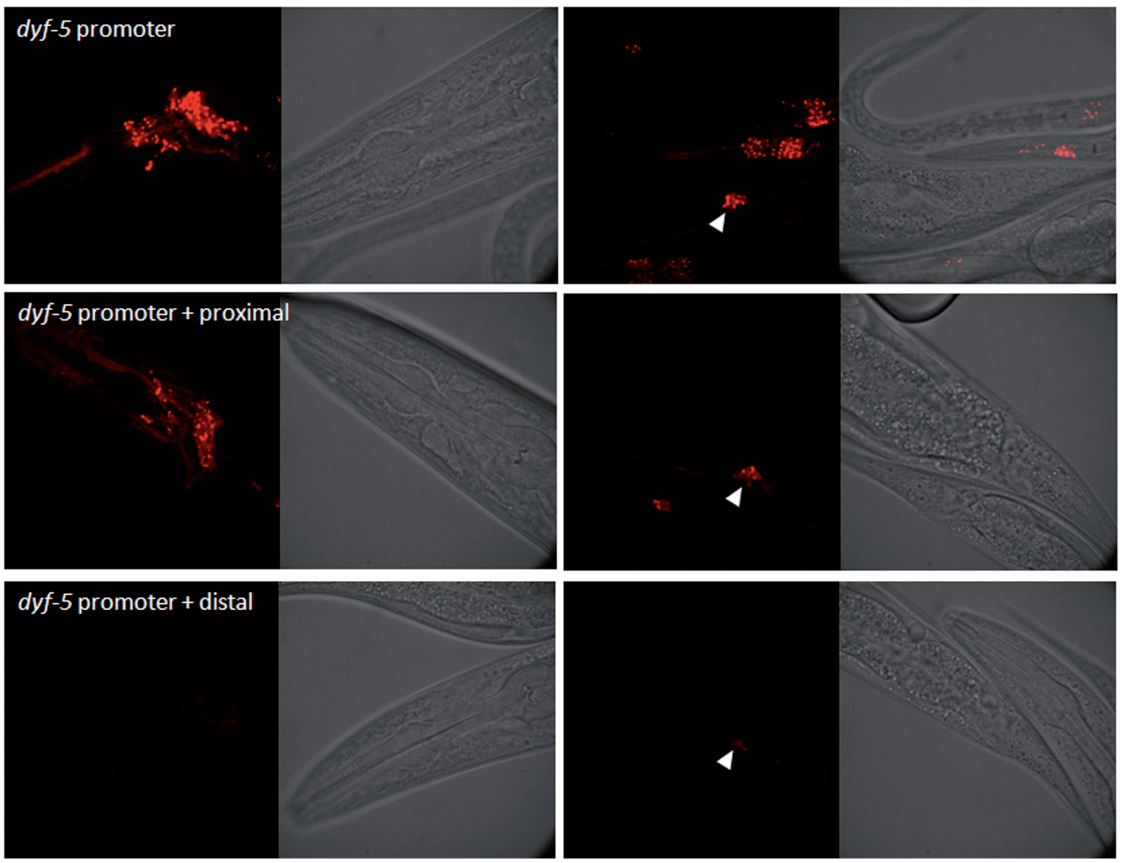

Figure 4. The expression pattern driven by *dyf-5* promoter replacing the endogenous X-box motif with either the proximal motif or the distal motif. Proximal motif is able to drive normal expression while distal motif is unable to. The white arrows show the location of PHA and PHB neurons. Exposure time = 3 s.

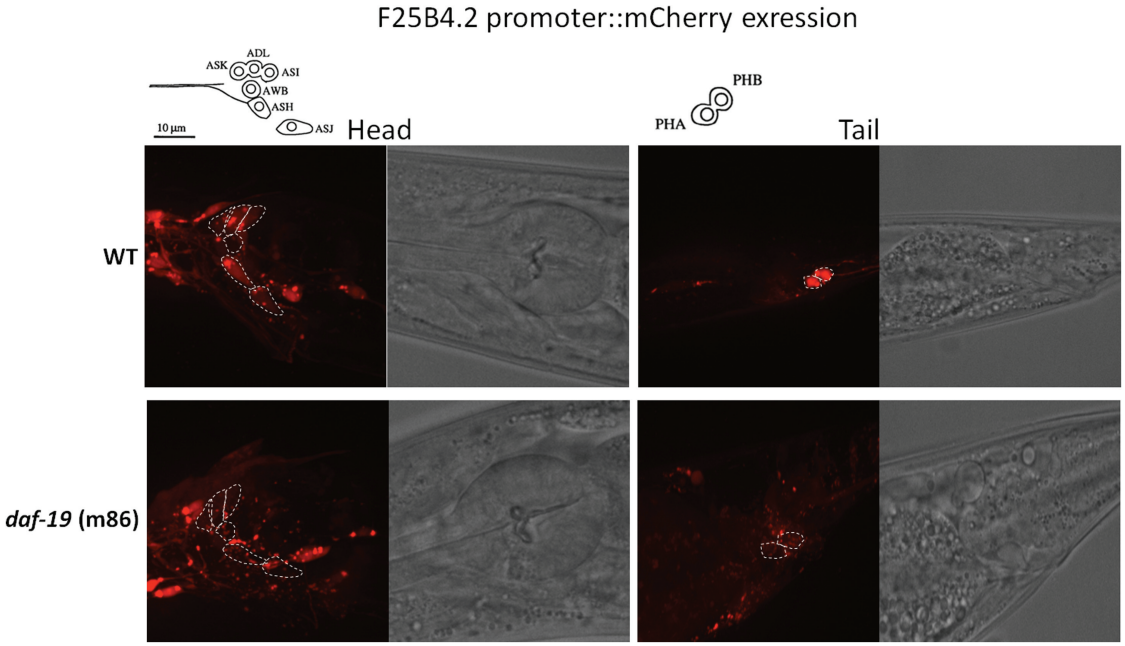

Figure 5. The head and tail expression patterns of F25B4.2 3 kb upstream region fused to mCherry in either wild-type (WT) strain or *daf-19(m86)* strain. White dashed lines outline the ciliated neurons that dye fill. Neurons that dye fill in the head include ASK, ADL, ASI, AWB, ASH and ASJ; neurons that dye fill in the tail include PHA and PHB. Because *daf-19* worms do not dye fill, the white outlines are the supposed location of these neurons. The expressions in these neurons are abolished in *daf-19(m86)* background. The schematic of the ciliated neurons that dye fill is shown on top. The schematic is adapted from (65). Exposure time = 3 s.

neurons (outlined by white dash lines in Figure 5). Expression in AWB was not found. Additional expression was also observed in muscle cells during larval stages but not in adults. Similar expression pattern for this gene was observed previously in *C. elegans* injected with extra-chromosomal array that contained GFP reporter driven by the same putative promoter sequence (38). To confirm whether the expression pattern indicated by mCherry is dependent on DAF-19, we crossed the strain with the mCherry reporter construct to a *daf-19* mutant strain (*m86*). In the absence of DAF-19, the expression in ciliated neurons that dye-fills in the head and tail is abolished (Figure 5). The fact that mCherry signals remain in other neurons suggests that additional *cis*- and *trans*- acting factors exist to regulate F25B4.2 expression pattern. Nevertheless, our data here shows that F25B4.2 expression in many ciliated neurons in *C. elegans* is dependent on DAF-19.

Both proximal and distal motifs function in regulating F25B4.2

To test whether these two motifs are functional in their endogenous environment, we engineered three additional promoter fusion constructs with (i) only the proximal motif removed, (ii) only the distal motif removed and (iii) both the proximal and the distal motifs removed (Figure 6). If these motifs are functional in promoting gene expression, we would expect the expression pattern in ciliated neurons to be abolished. These constructs were also injected and integrated using the MosSCI method (61). In the strain carrying the proximal deletion construct (JNC21), we observed that many amphid neurons as well as phasmid neurons lost mCherry expression (compare Figure 6a and b; i and j). Using dye-filling with DiO, we observed specifically that ASK, ASI and ASJ neurons no longer show expression while ADL and ASH neurons retained expression. In the strain carrying the distal deletion construct (JNC22), we were surprised that it did not abolish any expression but instead enhanced expression (compare Figure 6a and c; i and k). By reducing the exposure time from 3 s to 800 ms (~4-fold), we were able to capture the expression intensity at a comparable level to that of JNC20 (the strain carrying the wild-type promoter). In the strain carrying construct with both motifs removed (JNC29), we observed similar pattern and intensity as the JNC21 where many ciliated neurons no longer show mCherry expression (Figure 6). Again, dye-filing with DiO reveals that ASK, ASI, ASJ neurons do not show mCherry expression anymore while ADL and ASH neurons retained expression. Since we used a 3-kb promoter region in our constructs, it is possible that additional X-box motifs outside of the 500-bp region are responsible for the expression remaining in ADL and ASH neurons. Searching further upstream revealed a 13-bp motif at −1092 with a weak match to X-box motifs consensus (results not shown), which can be a *bona fide* X-box motif. Taken together, our results suggest proximal motif but not the distal motif is responsible for driving F25B4.2 expression in ciliated neurons. However the distal motif may have a repressive role in modulating the expression level of this gene.

In order to show whether these motifs function together with DAF-19, we have crossed JNC21, JNC22 and JNC29 to the *daf-19(m86)* mutant background. If the X-box

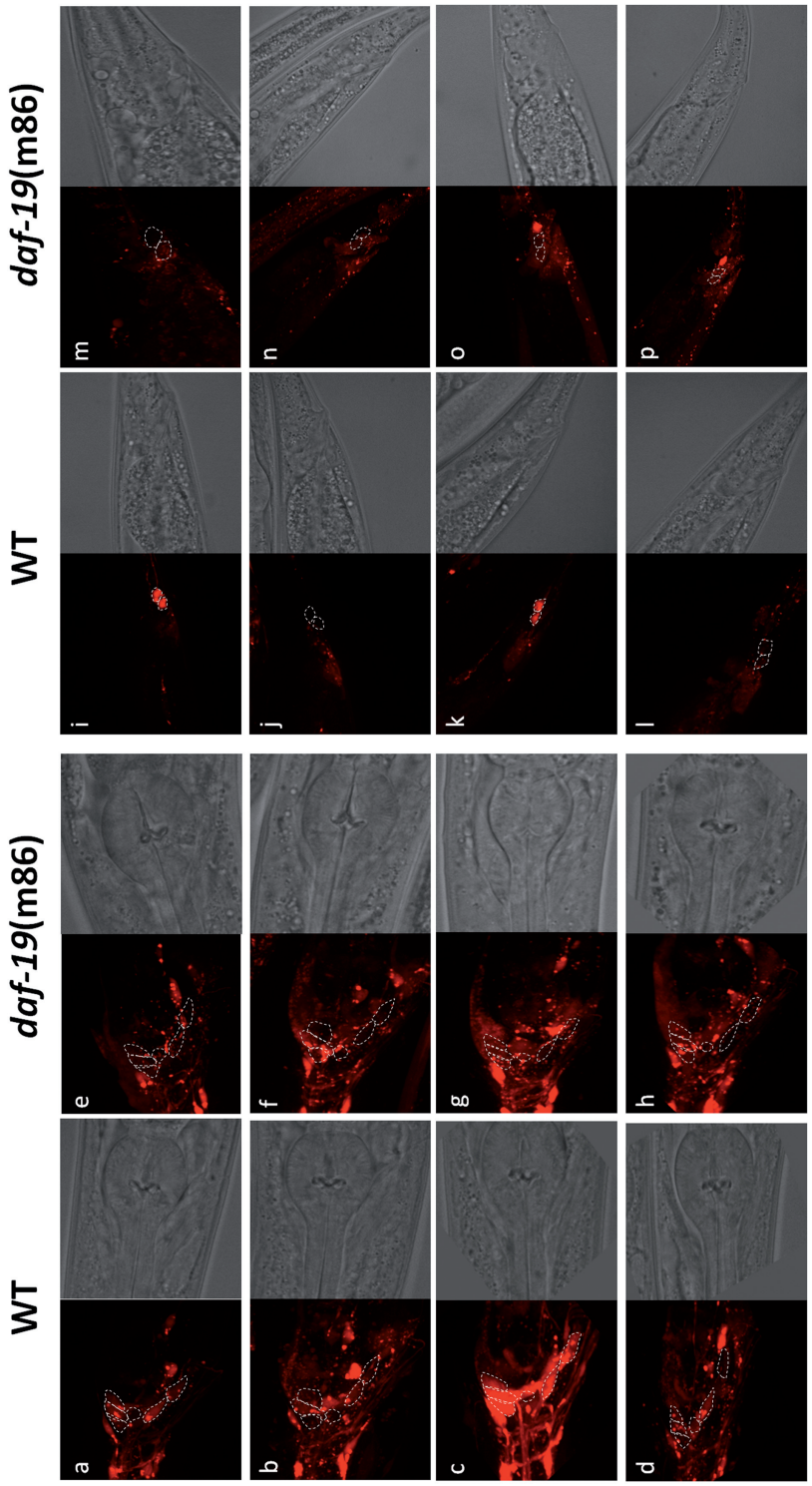

Figure 6. The expression of different deletion constructs in either wild-type or *daf-19(m86)* backgrounds. Proximal deletion construct removes the 15 bp putative X-box at −140; distal deletion construct removes the 15-bp putative X-box at −190; double deletion construct removes both putative X-box motifs. Other than that, all sequences remain the same. White dashed lines outline the ciliated neurons that dye fill. The outlines for strains in *daf-19* background are supposed locations. (**a**–**g**) show expression in the head and (**i**–**p**) show expression in the tail. Exposure time = 3 s.

motifs are functional with DAF-19, we expect these constructs in the *daf-19* mutant background would show similar pattern to what was observed in Figure 5 where many ciliated neurons no longer show mCherry expression in *daf-19* mutant worms. As expected, we observed nearly identical expression pattern across all constructs in *daf-19* mutant strain where many ciliated neurons in both the head and tail have abolished expression (compare Figure 6e to h; m to p). The difference is especially striking for the distal deletion construct where an elevated expression level in wild-type background dropped to very low expression in *daf-19* mutant (Figure 6c and g). The observation here further suggests that proximal motif is the main driving force for expression by interacting with DAF-19.

DISCUSSION

Ciliopathy is an emerging human genetic disorder caused by malformation of cilia that leads to many clinical hallmarks including obesity, polydactyly and retinal degeneration. The first link between RFX TFs and ciliary genes together with cilia development was made by examining sensory defect in *C. elegans*, which lead to the cloning of *daf-19* (4), the only RFX TF gene in *C. elegans*. In the 10 years that followed this discovery, studies on *Chlamydomonas reinhardtii* and *C. elegans* greatly facilitated further ciliopathy research in mammals. *Caenorhabditis elegans*, in particular, has been used as a platform for identifying human BBS3, 5, 7 and 8 (48,50,66). Evidence suggests that the known 14 human BBS genes only constitute 25–50% of the ciliopathy cases (67). Therefore, additional target genes remain to be discovered. More importantly, mechanisms underlying the transcriptional regulation of these target genes by RFX TFs are not known. Such knowledge can definitely facilitate further search for RFX TF target genes.

DAF-19 target genes identified in *C. elegans* to date have been demonstrated to be regulated by X-box motifs. In this study, through bioinformatics and comparative genomics searches, we have found genes that have two putative X-box motifs residing in 500-bp upstream region. To ascertain the contribution of each individual X-box motif in gene expression and the possible interaction of two X-box motifs, we extensively examined the two 15-bp motifs in the genomic region upstream of F25B4.2. We show that F25B4.2 is regulated by both X-box motifs. F25B4.2 is likely orthologous to human Pellino gene. Pellino proteins are E3 ligases known to participate in balancing inflammatory response (56). Pellino proteins interact with IRAK and mediate NFkB nuclear translocation to promote activation of proinflammatory genes (57,58). Pellino1 is suggested to play a part in TGF-β pathways to promote anti-inflammatory response preventing hyperactivation of inflammatory response (59,60). However, Pellino is not currently known to have any role in cilia development or cilia maintenance in human or in any other organisms.

The proximal motif identified in this study can be seen as a 'strong' motif that has higher sequence conservation and drives gene expression while the distal motif can be seen as a 'weak' motif that do not drive gene expression as well but may function in expression level regulation. While a repressive X-box motif was never reported in *C. elegans*, there are ample of repressive X-box motifs in other systems. For example, RFX TF in *S. cerevisiae* was found to exhibit similar characteristic. *S. cerevisiae* RFX negatively regulates many ribonucleotide reductase genes (e.g. RNR2, RNR3 and RNR4) through a combination of strong X-box motifs and weak X-box motifs (24). Removing the weak X-box motifs only show slight expression increase (1.4–1.7 fold) and removing the strong X-box motif increase the expression level by 5-fold (24). However, simultaneous removal of all motifs elevates the expression level by 17-fold (24). Other TFs have also been found to bind to multiple regulatory elements in order to regulate gene expression to a specific level. For instance, PurR in *Bacillus subtilis* binds to two PurBox motifs for higher affinity (68). *Drosophila* gap gene hunchback (hb) is regulated by multiple bicoid (bcd) binding sites (69). Multiple NFkB binding sites work synergistically to regulate US3 gene of human cytomegalovirus (70).

In a similar fashion, human RFX1 and RFX3 repress MAP1A expression in non-neuronal cells by binding to two X-box motifs in the first exon (25). The strong and weak X-box motifs in yeast and humans may work synergistically; however, our results here suggest that X-box motifs in F25B4.2 work antagonistically. We propose that while a single strong X-box is sufficient to drive gene expression, additional weak X-box motifs could fine-tune the expression level appropriate for the gene. The molecular mechanism of the fine-tuning remains to be further explored. A simple model is that the weak motif servers as a dominant negative competitor against the strong motif. In other words, the weak site servers as a 'transcription factor sponge' that absorbs transcription factor in the local molecular niche so that the strong site has limited supply of the transcription factor.

In conclusion, this project represents an important step toward understanding RFX regulatory mechanism in animal genome, which will potentially help identifying functional X-box motifs as well as RFX target genes in humans.

ACKNOWLEDGEMENT

We thank Erik M. Jorgensen for plasmids and strains and Harald Hutter for discussion of results. We thank Tao Luo, Elaheh Kahrobaee Khorassani and Behnaz Kahrobaee Khorassani for technical assistance. Reagents for MosSCI were generous gifts from E.M. Jorgensen.

FUNDING

Discovery Grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada (to N.C. and D.L.B.). J.S.C.C. is supported by a NSERC Postgraduate Scholarship. D.L.B. is a Tier 1 Canada

Research Chair (CRC). N.C. is a Michael Smith Foundation for Health Research (MSFHR) Scholar and a Canadian Institutes of Health Research (CIHR) New Investigator. Funding for open access charge: NSERC Canada.

Conflict of interest statement. None declared.

REFERENCES

1. Emery,P., Durand,B., Mach,B. and Reith,W. (1996) RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Res.*, **24**, 803–807.
2. Chu,J.S., Baillie,D.L. and Chen,N. (2010) Convergent evolution of RFX transcription factors and ciliary genes predated the origin of metazoans. *BMC Evol. Biol.*, **10**, 130.
3. Gajiwala,K.S., Chen,H., Cornille,F., Roques,B.P., Reith,W., Mach,B. and Burley,S.K. (2000) Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature*, **403**, 916–921.
4. Swoboda,P., Adler,H.T. and Thomas,J.H. (2000) The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. *Mol. Cell*, **5**, 411–421.
5. Durand,B., Vandaele,C., Spencer,D., Pantalacci,S. and Couble,P. (2000) Cloning and characterization of dRFX, the Drosophila member of the RFX family of transcription factors. *Gene*, **246**, 285–293.
6. Otsuki,K., Hayashi,Y., Kato,M., Yoshida,H. and Yamaguchi,M. (2004) Characterization of dRFX2, a novel RFX family protein in Drosophila. *Nucleic Acids Res.*, **32**, 5636–5648.
7. Hoskins,R.A., Carlson,J.W., Kennedy,C., Acevedo,D., Evans-Holm,M., Frise,E., Wan,K.H., Park,S., Mendez-Lago,M., Rossi,F. *et al.* (2007) Sequence finishing and mapping of Drosophila melanogaster heterochromatin. *Science*, **316**, 1625–1628.
8. Aftab,S., Semenec,L., Chu,J.S. and Chen,N. (2008) Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evol. Biol.*, **8**, 226.
9. Bonnafe,E., Touka,M., AitLounis,A., Baas,D., Barras,E., Ucla,C., Moreau,A., Flamant,F., Dubruille,R., Couble,P. *et al.* (2004) The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. *Mol. Cell. Biol.*, **24**, 4417–4427.
10. Baas,D., Meiniel,A., Benadiba,C., Bonnafe,E., Meiniel,O., Reith,W. and Durand,B. (2006) A deficiency in RFX3 causes hydrocephalus associated with abnormal differentiation of ependymal cells. *Eur. J. Neurosci.*, **24**, 1020–1030.
11. Ait-Lounis,A., Baas,D., Barras,E., Benadiba,C., Charollais,A., Nlend Nlend,R., Liegeois,D., Meda,P., Durand,B. and Reith,W. (2007) Novel function of the ciliogenic transcription factor RFX3 in development of the endocrine pancreas. *Diabetes*, **56**, 950–959.
12. Ashique,A.M., Choe,Y., Karlen,M., May,S.R., Phamluong,K., Solloway,M.J., Ericson,J. and Peterson,A.S. (2009) The Rfx4 transcription factor modulates Shh signaling by regional control of ciliogenesis. *Sci. Signal.*, **2**, ra70.
13. Praetorius,H.A. and Spring,K.R. (2005) A physiological view of the primary cilium. *Ann. Rev. Physiol.*, **67**, 515–529.
14. Haycraft,C.J., Schafer,J.C., Zhang,Q., Taulman,P.D. and Yoder,B.K. (2003) Identification of CHE-13, a novel intraflagellar transport protein required for cilia formation. *Exp. Cell Res.*, **284**, 251–263.
15. Badano,J.L., Mitsuma,N., Beales,P.L. and Katsanis,N. (2006) The ciliopathies: an emerging class of human genetic disorders. *Ann. Rev. Genomics Hum. Genet.*, **7**, 125–148.
16. Garvie,C.W., Stagno,J.R., Reid,S., Singh,A., Harrington,E. and Boss,J.M. (2007) Characterization of the RFX complex and the RFX5(L66A) mutant: implications for the regulation of MHC class II gene expression. *Biochemistry*, **46**, 1597–1611.
17. Soyer,J., Flasse,L., Raffelsberger,W., Beucher,A., Orvain,C., Peers,B., Ravassard,P., Vermot,J., Voz,M.L., Mellitzer,G. *et al.* (2010) Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development. *Development*, **137**, 203–212.
18. Smith,S.B., Qu,H.Q., Taleb,N., Kishimoto,N.Y., Scheel,D.W., Lu,Y., Patch,A.M., Grabs,R., Wang,J., Lynn,F.C. *et al.* (2010) Rfx6 directs islet formation and insulin production in mice and humans. *Nature*, **463**, 775–780.
19. Emery,P., Strubin,M., Hofmann,K., Bucher,P., Mach,B. and Reith,W. (1996) A consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity. *Mol. Cell. Biol.*, **16**, 4486–4494.
20. Laurencon,A., Dubruille,R., Efimenko,E., Grenier,G., Bissett,R., Cortier,E., Rolland,V., Swoboda,P. and Durand,B. (2007) Identification of novel regulatory factor X (RFX) target genes by comparative genomics in Drosophila species. *Genome Biol.*, **8**, R195.
21. Efimenko,E., Bubb,K., Mak,H.Y., Holzman,T., Leroux,M.R., Ruvkun,G., Thomas,J.H. and Swoboda,P. (2005) Analysis of xbx genes in C. elegans. *Development*, **132**, 1923–1934.
22. Blacque,O.E., Perens,E.A., Boroevich,K.A., Inglis,P.N., Li,C., Warner,A., Khattra,J., Holt,R.A., Ou,G., Mah,A.K. *et al.* (2005) Functional genomics of the cilium, a sensory organelle. *Curr. Biol.*, **15**, 935–941.
23. Chen,N., Mah,A., Blacque,O.E., Chu,J., Phgora,K., Bakhoum,M.W., Newbury,C.R., Khattra,J., Chan,S., Go,A. *et al.* (2006) Identification of ciliary and ciliopathy genes in Caenorhabditis elegans through comparative genomics. *Genome Biol.*, **7**, R126.
24. Huang,M., Zhou,Z. and Elledge,S.J. (1998) The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. *Cell*, **94**, 595–605.
25. Nakayama,A., Murakami,H., Maeyama,N., Yamashiro,N., Sakakibara,A., Mori,N. and Takahashi,M. (2003) Role for RFX transcription factors in non-neuronal cell-specific inactivation of the microtubule-associated protein MAP1A promoter. *J. Biol. Chem.*, **278**, 233–240.
26. Brenner,S. (1974) The genetics of Caenorhabditis elegans. *Genetics*, **77**, 71–94.
27. She,R., Chu,J.S., Uyar,B., Wang,J. and Wang,K. (2011) genBlastG: using BLAST searches to build homologous gene models. *Bioinformatics*, **27**, 2141–2143.
28. Stein,L.D., Mungall,C., Shu,S., Caudy,M., Mangone,M., Day,A., Nickerson,E., Stajich,J.E., Harris,T.W., Arva,A. *et al.* (2002) The generic genome browser: a building block for a model organism system database. *Genome Res.*, **12**, 1599–1610.
29. Remm,M., Storm,C.E. and Sonnhammer,E.L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.*, **314**, 1041–1052.
30. Liefooghe,A., Touzet,H. and Varre,J.-S. (2006) Large scale matching for Position Weight Matrices. *Lect. Notes Comput. Sci.*, **4009**, 401–412.
31. Hobert,O. (2002) PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. *BioTechniques*, **32**, 728–730.
32. Cell identification in C. elegans. 2005. Altun, Z.F. and Hall, D.H. In WormAtlas. http://www.wormatlas.org/cellID.html.
33. Ye,S., Dhillon,S., Ke,X., Collins,A.R. and Day,I.N. (2001) An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.*, **29**, E88.
34. Perkins,L.A., Hedgecock,E.M., Thomson,J.N. and Culotti,J.G. (1986) Mutant sensory cilia in the nematode Caenorhabditis elegans. *Dev. Biol.*, **117**, 456–487.
35. Stein,L.D., Bao,Z., Blasiar,D., Blumenthal,T., Brent,M.R., Chen,N., Chinwalla,A., Clarke,L., Clee,C., Coghlan,A. *et al.* (2003) The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. *PLoS Biol.*, **1**, E45.
36. Haycraft,C.J., Swoboda,P., Taulman,P.D., Thomas,J.H. and Yoder,B.K. (2001) The C. elegans homolog of the murine cystic kidney disease gene Tg737 functions in a ciliogenic pathway and is disrupted in osm-5 mutant worms. *Development*, **128**, 1493–1505.
37. Qin,H., Rosenbaum,J.L. and Barr,M.M. (2001) An autosomal recessive polycystic kidney disease gene homolog is involved in intraflagellar transport in C. elegans ciliated sensory neurons. *Curr. Biol.*, **11**, 457–461.
38. Hunt-Newbury,R., Viveiros,R., Johnsen,R., Mah,A., Anastas,D., Fang,L., Halfnight,E., Lee,D., Lin,J., Lorch,A. *et al.* (2007)

High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. *PLoS Biol.*, **5**, e237.
39. McKay,S.J., Johnsen,R., Khattra,J., Asano,J., Baillie,D.L., Chan,S., Dube,N., Fang,L., Goszczynski,B., Ha,E. *et al.* (2003) Gene expression profiling of cells, tissues, and developmental stages of the nematode C. elegans. *Cold Spring Harb. Symp. Quant. Biol.*, **68**, 159–169.
40. Schafer,J.C., Haycraft,C.J., Thomas,J.H., Yoder,B.K. and Swoboda,P. (2003) XBX-1 encodes a dynein light intermediate chain required for retrograde intraflagellar transport and cilia assembly in Caenorhabditis elegans. *Mol. Biol. Cell*, **14**, 2057–2070.
41. Winkelbauer,M.E., Schafer,J.C., Haycraft,C.J., Swoboda,P. and Yoder,B.K. (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.
42. Efimenko,E., Blacque,O.E., Ou,G., Haycraft,C.J., Yoder,B.K., Scholey,J.M., Leroux,M.R. and Swoboda,P. (2006) Caenorhabditis elegans DYF-2, an orthologue of human WDR19, is a component of the intraflagellar transport machinery in sensory cilia. *Mol. Biol. Cell*, **17**, 4801–4811.
43. Collet,J., Spike,C.A., Lundquist,E.A., Shaw,J.E. and Herman,R.K. (1998) Analysis of osm-6, a gene that affects sensory cilium structure and sensory neuron function in Caenorhabditis elegans. *Genetics*, **148**, 187–200.
44. Ou,G., Qin,H., Rosenbaum,J.L. and Scholey,J.M. (2005) The PKD protein qilin undergoes intraflagellar transport. *Curr. Biol.*, **15**, R410–R411.
45. Murayama,T., Toh,Y., Ohshima,Y. and Koga,M. (2005) The dyf-3 gene encodes a novel protein required for sensory cilium formation in Caenorhabditis elegans. *J. Mol. Biol.*, **346**, 677–687.
46. Fujiwara,M., Ishihara,T. and Katsura,I. (1999) A novel WD40 protein, CHE-2, acts cell-autonomously in the formation of C. elegans sensory cilia. *Development*, **126**, 4839–4848.
47. Bell,L.R., Stone,S., Yochem,J., Shaw,J.E. and Herman,R.K. (2006) The molecular identities of the Caenorhabditis elegans intraflagellar transport genes dyf-6, daf-10 and osm-1. *Genetics*, **173**, 1275–1286.
48. Blacque,O.E., Reardon,M.J., Li,C., McCarthy,J., Mahjoub,M.R., Ansley,S.J., Badano,J.L., Mah,A.K., Beales,P.L., Davidson,W.S. *et al.* (2004) Loss of C. elegans BBS-7 and BBS-8 protein function results in cilia defects and compromised intraflagellar transport. *Genes Dev.*, **18**, 1630–1642.
49. Ansley,S.J., Badano,J.L., Blacque,O.E., Hill,J., Hoskins,B.E., Leitch,C.C., Kim,J.C., Ross,A.J., Eichers,E.R., Teslovich,T.M. *et al.* (2003) Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. *Nature*, **425**, 628–633.
50. Li,J.B., Gerdes,J.M., Haycraft,C.J., Fan,Y., Teslovich,T.M., May-Simera,H., Li,H., Blacque,O.E., Li,L., Leitch,C.C. *et al.* (2004) Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. *Cell*, **117**, 541–552.
51. Mak,H.Y., Nelson,L.S., Basson,M., Johnson,C.D. and Ruvkun,G. (2006) Polygenic control of Caenorhabditis elegans fat storage. *Nat. Genet.*, **38**, 363–368.
52. Bacaj,T., Lu,Y. and Shaham,S. (2008) The conserved proteins CHE-12 and DYF-11 are required for sensory cilium function in Caenorhabditis elegans. *Genetics*, **178**, 989–1002.
53. Burghoorn,J., Dekkers,M.P., Rademakers,S., de Jong,T., Willemsen,R. and Jansen,G. (2007) Mutation of the MAP kinase DYF-5 affects docking and undocking of kinesin-2 motors and reduces their speed in the cilia of Caenorhabditis elegans. *Proc. Natl Acad. Sci. USA*, **104**, 7157–7162.
54. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
55. Nicholas,K.B., Nicholas,H.B.J. and Deerfield,D.W.I. (1997) GeneDoc: analysis and visualization of genetic variation. *EMBNET.NEWS*, **4**.
56. Butler,M.P., Hanly,J.A. and Moynagh,P.N. (2007) Kinase-active interleukin-1 receptor-associated kinases promote polyubiquitination and degradation of the Pellino family: direct evidence for PELLINO proteins being ubiquitin-protein isopeptide ligases. *J. Biol. Chem.*, **282**, 29729–29737.
57. Rich,T., Allen,R.L., Lucas,A.M., Stewart,A. and Trowsdale,J. (2000) Pellino-related sequences from Caenorhabditis elegans and Homo sapiens. *Immunogenetics*, **52**, 145–149.
58. Strelow,A., Kollewe,C. and Wesche,H. (2003) Characterization of Pellino2, a substrate of IRAK1 and IRAK4. *FEBS Lett.*, **547**, 157–161.
59. Chang,M., Jin,W. and Sun,S.C. (2009) Peli1 facilitates TRIF-dependent Toll-like receptor signaling and proinflammatory cytokine production. *Nat. Immunol.*, **10**, 1089–1095.
60. Choi,K.C., Lee,Y.S., Lim,S., Choi,H.K., Lee,C.H., Lee,E.K., Hong,S., Kim,I.H., Kim,S.J. and Park,S.H. (2006) Smad6 negatively regulates interleukin 1-receptor-Toll-like receptor signaling through direct interaction with the adaptor Pellino-1. *Nat. Immunol.*, **7**, 1057–1065.
61. Frokjaer-Jensen,C., Davis,M.W., Hopkins,C.E., Newman,B.J., Thummel,J.M., Olesen,S.P., Grunnet,M. and Jorgensen,E.M. (2008) Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat. Genet.*, **40**, 1375–1383.
62. Benjamin,H.W. and Kleckner,N. (1992) Excision of Tn10 from the donor site during transposition occurs by flush double-strand cleavages at the transposon termini. *Proc. Natl Acad. Sci. USA*, **89**, 4648–4652.
63. Lampe,D.J., Churchill,M.E. and Robertson,H.M. (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J.*, **15**, 5470–5479.
64. van Luenen,H.G., Colloms,S.D. and Plasterk,R.H. (1994) The mechanism of transposition of Tc3 in C. elegans. *Cell*, **79**, 293–301.
65. Starich,T.A., Herman,R.K., Kari,C.K., Yeh,W.H., Schackwitz,W.S., Schuyler,M.W., Collet,J., Thomas,J.H. and Riddle,D.L. (1995) Mutations affecting the chemosensory neurons of Caenorhabditis elegans. *Genetics*, **139**, 171–188.
66. Fan,Y., Esmail,M.A., Ansley,S.J., Blacque,O.E., Boroevich,K., Ross,A.J., Moore,S.J., Badano,J.L., May-Simera,H., Compton,D.S. *et al.* (2004) Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet-Biedl syndrome. *Nat. Genet.*, **36**, 989–993.
67. Yang,Z., Yang,Y., Zhao,P., Chen,K., Chen,B., Lin,Y., Guo,F., Chen,Y., Liu,X., Lu,F. *et al.* (2008) A novel mutation in BBS7 gene causes Bardet-Biedl syndrome in a Chinese family. *Mol. Vis.*, **14**, 2304–2308.
68. Bera,A.K., Zhu,J., Zalkin,H. and Smith,J.L. (2003) Functional dissection of the Bacillus subtilis pur operator site. *J. Bacteriol.*, **185**, 4099–4109.
69. Driever,W. and Nusslein-Volhard,C. (1989) The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. *Nature*, **337**, 138–143.
70. Chan,Y.J., Tseng,W.P. and Hayward,G.S. (1996) Two distinct upstream regulatory domains containing multicopy cellular transcription factor binding sites provide basal repression and inducible enhancer characteristics to the immediate-early IES (US3) promoter from human cytomegalovirus. *J. Virol.*, **70**, 5312–5328.