Analysis of *xbx* genes in *C. elegans*

Evgeni Efimenko[1], Kerry Bubb[2], Ho Yi Mak[4], Ted Holzman[3], Michel R. Leroux[5], Gary Ruvkun[4], James H. Thomas[2] and Peter Swoboda[1,*]

[1]Karolinska Institute, Department of Biosciences and Södertörn University College, Section of Natural Sciences, S-14189 Huddinge, Sweden
[2]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
[3]Department of Microbiology, University of Washington, Seattle, WA 98195, USA
[4]Department of Molecular Biology, Massachusetts General Hospital and Department of Genetics, Harvard Medical School, Boston, MA 02114, USA
[5]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada
*Author for correspondence (e-mail: peter.swoboda@biosci.ki.se)

Accepted 10 February 2005

Development 132, 1923-1934
Published by The Company of Biologists 2005
doi:10.1242/dev.01775

Summary

Cilia and flagella are widespread eukaryotic subcellular components that are conserved from green algae to mammals. In different organisms they function in cell motility, movement of extracellular fluids and sensory reception. While the function and structural description of cilia and flagella are well established, there are many questions that remain unanswered. In particular, very little is known about the developmental mechanisms by which cilia are generated and shaped and how their components are assembled into functional machineries. To find genes involved in cilia development we used as a search tool a promoter motif, the X-box, which participates in the regulation of certain ciliary genes in the nematode *Caenorhabditis elegans*.

By using a genome search approach for X-box promoter motif-containing genes (*xbx* genes) we identified a list of about 750 *xbx* genes (candidates). This list comprises some already known ciliary genes as well as new genes, many of which we hypothesize to be important for cilium structure and function. We derived a *C. elegans* X-box consensus sequence by in vivo expression analysis. We found that *xbx* gene expression patterns were dependent on particular X-box nucleotide compositions and the distance from the respective gene start. We propose a model where DAF-19, the RFX-type transcription factor binding to the X-box, is responsible for the development of a ciliary module in *C. elegans*, which includes genes for cilium structure, transport machinery, receptors and other factors.

Key words: X-box, DAF-19, Ciliary genes, *C. elegans*

Introduction

Sensory behavior in higher animals depends on the correct recognition and processing of signals from the environment or from within the organism, and then on the appropriate reaction to those signals. The first step in this chain of events, signal recognition, is mediated by specialized cells of the nervous system, the sensory neurons. They create functional compartments for the localization and exposure of the signal reception and transduction machineries. There is considerable diversity in these compartments. Prominent examples are ciliated endings at the tip of dendrites of sensory neurons. Specialized mammalian sensory cilia include those in the photoreceptor cells in the eye, in the hair cells in the ear, and in olfactory neurons in the nose. The universal architecture of cilia and their close relatives, flagella, consists of a microtubular axonemal core enclosed by a membrane, exposed on the cell surface. A convergence of results from several systems has led to detailed structural descriptions of cilia and flagella (Perkins et al., 1986; Dutcher, 1995; Rosenbaum and Witman, 2002).

In the nematode *C. elegans* 60 of the 302 neurons of the hermaphrodite are ciliated sensory neurons (CSN) (Ward et al., 1975; White et al., 1986), forming many structurally distinct types of sensory cilia. Whereas many ciliary mutants are available in *C. elegans*, there is only one known gene mutation that completely eliminates all classes of sensory cilia and all functional components of cilium structure. This gene is *daf-19*, and it encodes the sole *C. elegans* member of the RFX-type transcription factors (Swoboda et al., 2000), found widely in the eukaryotic kingdom. All members of the RFX transcription factor family are characterized by the presence of a conserved DNA binding domain (DBD). The RFX-DBD binds to special motifs (X-boxes) in promoters of its target genes. Genes containing the X-box promoter motif are called *xbx* genes.

Apart from *C. elegans* DAF-19, the RFX family currently contains nine characterized members: five in mice and humans (RFX1-5) (Emery et al., 1996a), two in *Drosophila melanogaster* (dRFX1-2) (Dubruille et al., 2002; Otsuki et al., 2004), and one member each from *Schizosaccharomyces pombe* (Wu and McLeod, 1995) and *Saccharomyces cerevisiae* (Huang et al., 1998). The data obtained so far suggest diverse biological roles of RFX proteins. In yeasts they regulate some aspects of the cell cycle (Wu and McLeod, 1995; Huang et al., 1998). In humans RFX factors are involved in the transcriptional regulation of major histocompatibility complex class II genes (RFX5) (Reith and Mach, 2001) and in the

Development

modulation of Ras signaling in epithelial cells (RFX3) (Maijgren et al., 2004).

The finding of X-boxes in promoters of ciliary genes in *C. elegans* has revealed an important role of the RFX family in the regulation of ciliogenesis (Swoboda et al., 2000). Since then, the conservation of RFX-binding elements has been reported in several distantly related species. For example, some ciliary genes in *D. melanogaster* contain X-box-like sequences in their promoters (Avidor-Reiss et al., 2004). However, an experimental demonstration of an RFX-dependence for *Drosophila* ciliary gene candidates exists so far only for the *nompB* gene (Han et al., 2003). Recently, data about the possible RFX regulation of ciliogenesis in mammals were also obtained. *Rfx3*-deficient mice exhibit frequent left-right asymmetry defects, which are caused by ciliary abnormalities in mutant embryos (Bonnafe et al., 2004). Mouse RFX3 regulates the expression of *D2lic*, the mouse ortholog of the *C. elegans* ciliary gene *xbx-1*, but does not affect the expression of *Tg737*, the mouse ortholog of the *C. elegans* ciliary gene *osm-5*. These observations suggest that RFX regulation of ciliogenesis in higher organisms is more complicated, and different subtypes of RFX proteins may be restricted to particular components of ciliary structure.

In our current work we first concentrated on the isolation of genes important for cilium structure and function using a genome-wide X-box promoter motif search in the nematode species *C. elegans* and *C. briggsae*. In this computational approach we focused our efforts only on 5′ flanking regions of genes, since X-box motifs have previously been shown to be functional in those regions (Swoboda et al., 2000). Subsequently, we performed expression analyses of the group of positive *C. elegans* matches in wild-type and *daf-19* mutant backgrounds, together with X-box mutagenesis experiments. Results of these analyses established the X-box consensus for *C. elegans,* the approximate number of *xbx* genes in the *C. elegans* genome and assigned already known and newly found ciliary genes to specific structural and functional groups. Because the organization of *C. elegans* sensory cilia is very similar to sensory cilia in mammals, the results obtained with the *C. elegans* model will have general significance.

Materials and methods

Worm strains

Growth and culture of *C. elegans* strains were carried out following standard procedures (Brenner, 1974). The following strains were used for this study: wild type N2 Bristol; JT8651 *daf-19(m86)/mnC1; lin-15(n765ts)*; JT6924 *daf-19(m86); daf-12(sa204)*; JT204 *daf-12(sa204)*; RB773 *nud-1(ok552)*; RB819 *xbx-4(ok635)*; RB857 *xbx-6(ok852)*; NL2099 *rrf-3(pk1426)*; CB3323 *che-13(e1805)*; CB1033 *che-2(e1033)*. Extrachromosomal arrays were used for all GFP expression analyses. All strains used and strain construction details are available on request.

Promoter motif search algorithm and sequence analyses

The X-box motif search was performed primarily with a Perl-based algorithm that searches through a given genome sequence for all possible matches. The algorithm first finds all sequences that match a defined consensus. After that step, the main module of the program implements a cross-match file (P. Green, personal communication), which compares a 3 kb window downstream of each match to a file containing the DNA sequences for all predicted genes, and a file containing assembled ESTs. Cross-match parameters – ‘minmatch’ and ‘minscore’ were set to 40. All other parameters were kept at default values. Minimal and maximal distances from positive matches to predicted genes were set to a range of 0-1000 nucleotides. To obtain a copy of the algorithm, contact kbubb@u.washington.edu. Cross-match must be obtained separately (see www.phrap.com for access/download information).

Subsequent to genome analyses using Cross-match, we made use of another program, DNA Motif Searcher, which takes a set of user-definable X-box sequences to search for additional motif instances in the genome. The set of motifs is interpreted into a position-specific score matrix (PSSM). Using this PSSM, the program can then identify the closest matching occurrences of the motif based on a score cutoff, or it can identify the top number of occurrences of the motif. For download or for more detailed explanations about Motif Searcher, please contact jht@u.washington.edu.

Genome sequence information, EST files, gene predictions and identities for X-box searches were obtained from the following sources: *C. elegans* complete genome sequence, WS122 release (ftp://ftp.sanger.ac.uk/pub/wormbase/WS122); *C. briggsae* draft genome sequence, cb25.agp8 version (ftp://ftp.wormbase.org/pub/wormbase/briggsae).

Generation and analysis of expression constructs

GFP expression constructs were designed by inserting about 2 kb of promoter regions and the first several codons of a gene of interest into the GFP expression vector pPD95.77 (gift from A. Fire). PCR fragments of promoter regions were obtained from wild-type N2 genomic DNA and cloned into appropriate sites of pPD95.77. For some genes, the wild-type X-boxes within promoters were mutated by overlap extension mutagenesis, replacing X-box sequences with nonspecific nucleotides containing indicative restriction enzyme sites. To check for correct translational reading frames and promoter regions, junctions between vector and amplified inserts were verified by sequencing for all constructs. For the XBX-2::GFP translational fusion, the entire coding sequence of the gene with about 1 kb of promoter were fused to the pPD95.77 vector.

The following worm strains were used for injections and expression analyses: JT8651, JT6924 and JT204. The strain JT8651 *daf-19(m86)/mnC1; lin-15(n765ts)* served as the wild-type background, since *daf-19(m86)* is fully recessive. *daf-19* mutants are strongly Daf-c (dauer larva formation – constitutive) across the normal temperature range. Therefore, segregating dauers were recovered at 15°C to obtain a *daf-19* homozygous background. Alternatively, the strain JT6924 *daf-19(m86); daf-12(sa204)* was used as a *daf-19* mutant background. Worms of this genotype exhibit a Daf-d (dauer larva formation – defective) phenotype and do not require the recovery of dauers. In this case, JT204 *daf-12(sa204)* worms were used as a wild-type background with regard to *daf-19*.

Adult hermaphrodites were transformed using standard protocols (Mello et al., 1991). Constructs were injected typically at 10-100 ng/μl along with coinjection markers such as pRF4 (contains the dominant marker *rol-6(su1006)*) or pBLH98 (contains the wild-type *lin-15* gene to rescue *lin-15(n765ts)*).

Microscopy and imaging

GFP expression patterns were analyzed in stable transgenic lines at 1000× magnification by conventional fluorescence microscopy (Zeiss Axioplan 2). Expression patterns were examined in at least two independent transgenic lines at most developmental stages of the worm. Neuronal cell anatomies and identities followed published descriptions (Ward et al., 1975; White et al., 1986).

For the analysis of XBX-2::GFP movement properties, worms were mounted on agarose pads and anesthetized with 10 mM levamisol. Adult worms were analyzed with a Leica confocal imaging spectrophotometer TCS SP unit mounted on a Leica DMIRBE inverted microscope, and the obtained images were processed using

Development

Leica Confocal Software 2.5. Images were taken with a 63× objective and a 488 nm GFP filter. At least 40 stacked images were converted into an AVI file with a rate of two frames per second.

RNAi feeding experiments and fluorescent dye filling assays

RNA-mediated interference (RNAi) was performed according to standard methods (Timmons et al., 2001). PCR fragments for genes of interest were generated from N2 genomic DNA and cloned into the double T7 promoter-containing vector L4440 (gift from A. Fire). All constructs were transformed into HT115(DE3) bacterial cells and plated onto NGM plates with antibiotics and IPTG. L4-stage hermaphrodites were transferred to plates with induced bacteria and F_3 progenies were analyzed for possible phenotypes.

Fluorescent dye-filling assays were performed essentially as described previously (Starich et al., 1995) using the fluorescent dye DiI. Worm strains N2 and CB3323 were used as positive and negative controls, respectively. Stained adult hermaphrodites were analyzed at 1000× magnification by conventional fluorescence microscopy (Zeiss Axioplan 2).

Genetic characterization of *xbx* gene mutants

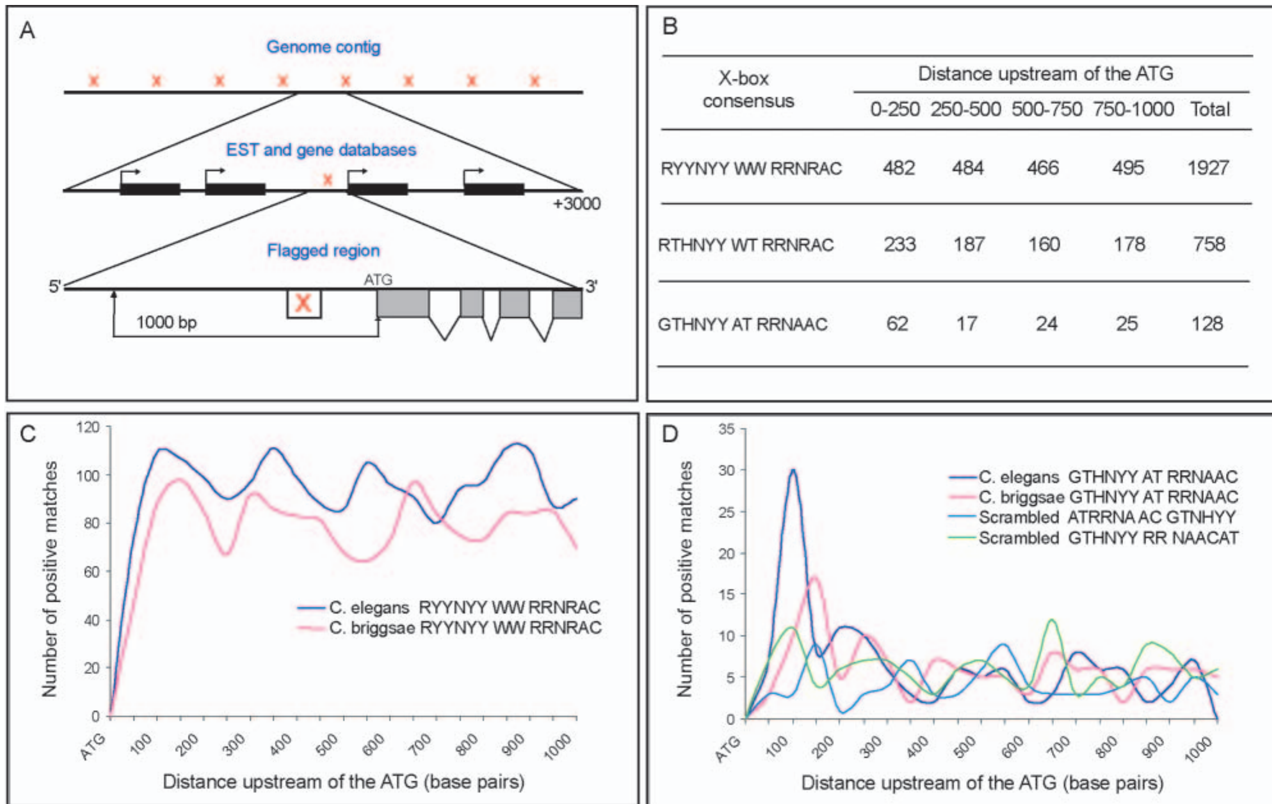
All deletion alleles analyzed in this study were generated by the *C. elegans* Gene Knockout Consortium (http://celeganskoconsortium.omrf.org/) using publicly available methodology (http://www.mutantfactory.ouhsc.edu/protocols.asp).

The original mutated strains RB819 *xbx-4(ok635)* IV and RB957 *xbx-6(ok852)* V were outcrossed three times with N2 and the following worm strains: JT7146 *egl-4(n478) unc-33(e204)* IV and DR108 *dpy-11(e224) unc-42(e270)* V, respectively. Outcrossed worms resulted in homozygous mutant strains containing the *xbx-4(ok635)* IV and *xbx-6(ok852)* V deletion alleles. These strains were then used as the basis for further analysis.

Results

Computational search for the X-box promoter motif

The discovery of X-boxes in promoter regions of certain ciliary genes (e.g. Swoboda et al., 2000; Haycraft et al., 2001; Haycraft et al., 2003; Fan et al., 2004) prompted us to analyze the whole *C. elegans* genome for the presence of these motifs. In order to perform this analysis of all *C. elegans* promoters we implemented an in-house searching algorithm – the X-box searcher, which searches for all possible matches to a defined motif sequence (Fig. 1A). For the initial search we used a 'relaxed' consensus (RYYNYY WW RRNRAC), that fits published mammalian X-box sequences (Emery et al., 1996b) and the first emerging *C. elegans* X-boxes from previously known ciliary genes (Swoboda et al., 2000). Using the X-box searcher with the relaxed consensus generates the highest number of output matches (1927) that were equally spread within 1000 bp of promoter regions (Fig. 1B,C). Through ongoing work with *xbx* gene candidates (see below) we obtained a 'refined' X-box consensus (GTHNYY AT



X-box consensus	Distance upstream of the ATG				
	0-250	250-500	500-750	750-1000	Total
RYYNYY WW RRNRAC	482	484	466	495	1927
RTHNYY WT RRNRAC	233	187	160	178	758
GTHNYY AT RRNAAC	62	17	24	25	128

Fig. 1. Computational search for the X-box motif. (A) Schematic of the X-box search algorithm. The algorithm finds all matches for a defined motif consensus and cross-matches them against a list of predicted genes or available ESTs. The search space upstream of predicted genes was set to 1000 bp. (B) The number of matches obtained with different consensus sequences (top to bottom: relaxed, average, refined) searching the *C. elegans* genome. (C) The 'relaxed' X-box consensus (RYYNYY WW RRNRAC), used for initial searches, generates the largest number of matches that spread equally within promoter regions. (D) The 'refined' X-box consensus (GTHNYY AT RRNAAC), obtained on the basis of in vivo expression analysis of *xbx* genes, shows a significant concentration of matches in the region of around 100 bp upstream of the ATG.

RRNAAC), which corresponds to most of the experimentally proven *xbx* genes. We further re-analyzed the *C. elegans* genome using the refined consensus and found a significant reduction in the number of matches (128). Unlike the relaxed consensus, most of the refined matches show clustering around the region of 100 bp upstream of the ATG. This clustering was not observed when the X-box search sequence was scrambled (Fig. 1B,D). The refined *xbx* gene list includes the majority of experimentally proven *xbx* genes, but not all of them. At the same time, the list of genes obtained with the relaxed X-box consensus contains too much ‘search noise’. According to our expression analysis (see below) we estimate that up to 90% of the relaxed matches can be found in promoter regions of non-ciliary genes or genes that are not specifically expressed in ciliated sensory neurons (CSN) (see Table S2 in supplementary material). Therefore, we additionally tried several different X-box sequences and found an ‘average’ consensus (RTHNYY WT RRNRAC). This consensus permitting ambiguities at three positions as compared to the refined consensus (A or G at position 1, A or T at position 7 and A or G at position 12) makes it possible to find all known X-boxes and reduces the search noise by about 2.5 times (758 matches) (Fig. 1B; see Table S1 in supplementary material). This consensus can be used for further analyses in *C. elegans*, as well as for X-box search efforts in other organisms.

The functional repertoire of the average candidate *xbx* gene list contains different molecular groups, including possible components of the ciliary structure and transport machinery (6%), transcription factors (5%), receptors (11%) and ion channels (1%). About 37% of the genes have no identified function and the rest of the list (40%) is composed of genes of various molecular identities (Fig. 2; see Table S1 in supplementary material).

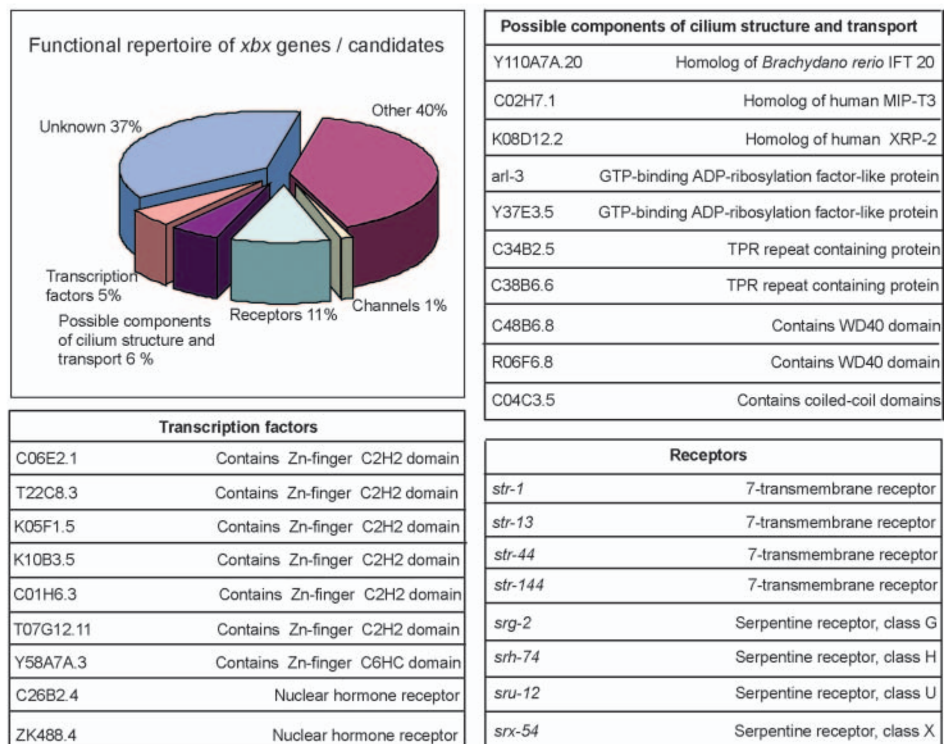
Our primary goal was to find new *xbx* and ciliary genes with high efficiency. Our X-box search might not be exhaustive, because the search algorithm is based on an originally small set of experimentally proven X-box sequences. To determine whether additional *xbx* gene candidates could be found in the *C. elegans* genome, we tried a different search approach, which is position-specific score matrix (PSSM) based (J.H.T., unpublished). Using the output data from the average list as a training set for the PSSM searcher we achieved almost the same list of *xbx* genes (data not shown), strengthening our overall strategy. However, we additionally obtained several prominent candidates that had a slightly different X-box motif than the average consensus and therefore were not found with the X-box searcher: for example *dlc-1* (GTTATT AT AAC*T*AC, which encodes a dynein light chain), C01B12.4 (GTTTCC AT AGC*T*AC, which encodes a predicted seven transmembrane receptor of the rhodopsin family).

Expression patterns of orthologous genes are often conserved. Because many orthologous transcription factors are also functionally conserved, one possible model to account for homologous gene expression patterns is conservation of specific binding sites within regulatory elements of orthologous genes (Ruvinsky et al., 2003). The nematodes *C. briggsae* and *C. elegans* are closely related species with very similar overall genome organizations (Stein et al., 2003). To find possible conservations of X-box regulatory elements between those two organisms we applied the *C. elegans* X-box search strategy to the *C. briggsae* genome. The number of obtained matches was slightly less than with *C. elegans*, probably because of the draft quality of the *C. briggsae* genome. Nevertheless, the profile of X-box distribution within promoter regions was similar to that of *C. elegans* (Fig. 1D).

In summary, using two different X-box promoter motif search approaches, X-box consensus sequences with varying degrees of refinement, together with cross-species comparisons and gene expression analysis (see below), we were able to identify a large number of bona fide *xbx* genes, a significant part of which we expect to also be ciliary- or CSN-specific genes.

Expression analysis of the *xbx* gene candidates

Our computational search has revealed a large, heterogeneous group of X-box matches (see Table S1 in supplementary



Possible components of cilium structure and transport	
Y110A7A.20	Homolog of *Brachydano rerio* IFT 20
C02H7.1	Homolog of human MIP-T3
K08D12.2	Homolog of human XRP-2
arl-3	GTP-binding ADP-ribosylation factor-like protein
Y37E3.5	GTP-binding ADP-ribosylation factor-like protein
C34B2.5	TPR repeat containing protein
C38B6.6	TPR repeat containing protein
C48B6.8	Contains WD40 domain
R06F6.8	Contains WD40 domain
C04C3.5	Contains coiled-coil domains

Transcription factors	
C06E2.1	Contains Zn-finger C2H2 domain
T22C8.3	Contains Zn-finger C2H2 domain
K05F1.5	Contains Zn-finger C2H2 domain
K10B3.5	Contains Zn-finger C2H2 domain
C01H6.3	Contains Zn-finger C2H2 domain
T07G12.11	Contains Zn-finger C2H2 domain
Y58A7A.3	Contains Zn-finger C6HC domain
C26B2.4	Nuclear hormone receptor
ZK488.4	Nuclear hormone receptor

Receptors	
str-1	7-transmembrane receptor
str-13	7-transmembrane receptor
str-44	7-transmembrane receptor
str-144	7-transmembrane receptor
srg-2	Serpentine receptor, class G
srh-74	Serpentine receptor, class H
sru-12	Serpentine receptor, class U
srx-54	Serpentine receptor, class X

Fig. 2. Functional repertoire of *xbx* genes. Functional repertoire of *xbx* genes and some representative members from different molecular groups, including possible components of cilium structure and transport machinery, transcription factors and receptors. The diagram represents data from the ‘average’ consensus list of *xbx* genes/candidates (758 members) (see Table S1 in supplementary material), including experimentally proven genes.

material). To determine the specificity of our computational search with regard to ciliary and CSN structure and function we initiated expression analyses for some of the genes. For this purpose we isolated a group of candidates with different nucleotide compositions of the X-box motifs, different positions within promoter regions and different proposed molecular functions of matching genes.

Our previous model associated the expression of *xbx* genes with CSN (Swoboda et al., 2000). We predicted that DAF-19 function, a particular X-box composition and its position within the promoter should be required for the expression of this group of genes. Therefore, expression patterns of selected genes were analyzed in both wild-type and *daf-19* mutant backgrounds. For some genes we also analyzed the actual X-box sequence by replacing it with nonspecific nucleotides within the respective expression construct.

Based on our obtained expression data (Table 1), we subdivide *xbx* genes into groups: (1) genes that are strongly regulated by DAF-19 and required for the general development of cilia; (2) genes that are partially regulated by DAF-19 and are probably required for more specific ciliary and/or CSN functions.

Genes of the first group are characterized by expression in many, most or all CSN. All of them are DAF-19 dependent and for some genes tested (through mutagenesis) we also demonstrated a dependence on the actual X-box sequence (Table 1).

Most genes from this group encode known participants of

Table 1. Expression analysis of *xbx* genes

C. elegans gene	X-box sequence (distance from ATG)	*C. briggsae* ortholog	X-box sequence (distance from ATG)	Expression patterns of *C. elegans* genes	Expression properties	References
Group 1						
*che-2** (F38G1.1)	GTTGTC AT GGTGAC (–130)	CBG13647	GTATCC AT GGCAAC (–182)	Many, most, all CSN	DD, XD	Fujiwara et al., 1999; Swoboda et al., 2000
*che-13** (F59C6.7)	GTTGCT AT AGCAAC (–75)	CBG02227	GTTTCC TT GACAAC (–85)	Many, most, all CSN	DD	Haycraft et.al., 2003
*osm-1** (T27B1.1)	GCTACC AT GGCAAC (–86)	CBG16355	GTTGCC AT GGACAC (–79)	Many, most, all CSN	DD, XD	Signor et al., 1999; Swoboda et al., 2000
*osm-5** (Y41G9A.1)	GTTACT AT GGCAAC (–116)	CBG02013	GTTGCC AG GGAAAC (–91)	Many, most, all CSN	DD, XD	Haycraft et.al., 2001
*osm-6** (R31.3)	GTTACC AT AGTAAC (–100)	CBG23329	X-box not found	Many, most, all CSN	DD, XD	Collet et al., 1998; Swoboda et al., 2000
*bbs-1** (Y105E8A.5)	GTTCCC AT AGCAAC (–99)	CBG08744	GTTGTT AT GGTAAC (–310)	Many, most, all CSN	DD	Ansley et al., 2003; current work
bbs-2 (F20D12.3)	GTATCC AT GGCAAC (–94)	CBG17712	ATATCC AT GGCAAC (–82)	Many, most, all CSN	DD, XD	Ansley et. al., 2003; current work
bbs-5 (R01H10.6)	GTCTCC AT GGCAAC (–66)	CBG23799	GTTACT AT GGCAAC (–69)	Many, most, all CSN	DD	Li et al., 2004
*bbs-7** (Y75B8A.12)	GTTGCC AT AGTAAC (–108)	CBG23043	GTTGCC AT GGTTAC (–138)	Many, most, all CSN	DD	Ansley et. al., 2003; current work
*bbs-8** (T25F10.5)	GTACCC AT GGCAAC (–84)	CBG19013	GTCTCT AT GGCAAC (–73)	Many, most, all CSN	DD	Ansley et. al., 2003; current work
*xbx-1** (F02D8.3)	GTTTCC AT GGTAAC (–79)	CBG11597	GTTTCC AT GGTTAC (–93)	Many, most, all CSN	DD, XD	Schafer et.al., 2003; current work
xbx-2 (D1009.5)	GTTGCC AT GACAAC (–78)	CBG00241	GTTTCC AT GGCTAC (–83)	Many, most, all CSN	DD, XD	Current work
Group 2						
xbx-3 (M04D8.6)	GTTGTC TT GGCAAC (–98)	CBG09908	GTTTCC AA GGAGAC (–128)	Amphids, phasmids	DD	Current work
xbx-4‡ (C23H5.3)	GTTGCC AT GACAAC (–82)	CBG10549	GTTGCC CT GGTGAC (–155)	Some CSN	DD	Current work
xbx-5 (T24A11.2)	GTCTCC AT GACAAC (–122)	CBG09228	GTCTCC AT GGCAAC (–142)	Some CSN	DD	Current work
xbx-6‡ (F40F9.1)	GTTTCC AT GGAAAC (–152)	CBG19349	GTATCC AT GGAAAC (–121)	Body wall muscles, pharyngeal muscles,ventral nerve cord, phasmids	DD (phasmids only)	Current work
xbx-7 (R148.1)	GTCACC AT AGGAAC (–70)	CBG22495	X-box not found	Labial neurons, some amphid neurons, phasmids	DD	Current work
nud-1 (F53A2.4)	GTATCC AT GAAAAC (–263)	CBG24281	X-box not found	Amphids, phasmids, vulva	DD (CSN only)	Dawe et al., 2001; current work
*che-11** (C27A7.4)	ATCTCC AT GGCAAC (–86)	CBG23392	GTATCC AT AGCAAC (–120)	Many, most, all CSN	DD (amphids and phasmids only)	Qin et al., 2001; current work
odr-4† (Y102E9.1)	ATCGTC AT CGTAAC (–164)	CBG16563	ATCGCC AT GGTTAC (–261)	10 amphid and 2 phasmid neurons	DD, XD	Dwyer et. al., 1998; current work
tub-1† (F10B5.4)	ATCTCC AT GACAAC (–183)	CBG00741	ATCACC AT GGCAAC (–232)	Many, most, all CSN	DD	Current work
nhr-44 (T19A5.4)	GTCTTC AT GGCAAC (–76)	CBG19141	X-box not found	ASK, other head neurons, other cell types	DD (in ASK)	Current work
Remaining genes analyzed						
F55D12.1	GTTACC AT AGTAAC (–234)	CBG08264	GTTGTC AT GACGAC (–252)	Glia, seam cells, vulva	DI	Current work
gpa-9‡ (F56H9.4)	GTTACC AT GGAAAC (–238)	Putative ortholog	not identified	ASJ, PHB, PVQ, pharyngeal muscle, spermatheca	DI	Jansen et al., 1999; current work
zag-1‡ (F28F9.1)	ATTGTC TA GGTAAC (–128)	CBG10736	X-box not found	Head and tail neurons	DI	Wacker et al., 2003; current work
F17A2.3	ACCGCC AA AGAAAC (–83)	Putative ortholog	not identified	Distal tip cells, ASJ, ASI	DI	Current work
aqp-2 (C01G6.1)	ACCACC TT GAAAAC (–115)	Putative ortholog	not identified	Excretory system, intestine, body wall muscle	DI	Current work

All genes analyzed were grouped according to their expression patterns and dependence on DAF-19. Upon mutation, genes marked with * cause abnormalities in general ciliary structure and function (e.g. Dyf+sensory phenotypes). Upon mutation, genes marked with † cause abnormalities in specific ciliary functions (e.g. Odr, but not Dyf phenotype). Upon mutation, genes marked with ‡, even though tested, have not yet shown abnormalities in general ciliary structure or specific ciliary functions. For genes not marked, mutants are presently not available. We tested gene expression for DAF-19 and X-box dependence, as indicated: DD, expression is DAF-19 dependent; XD, expression is X-box dependent; DI, expression is DAF-19 independent. CSN, ciliated sensory neurons. The expression patterns of some *C. elegans* genes were previously described, as referenced. The DAF-19 or X-box dependence for the genes *che-2*, *che-13*, *osm-1*, *osm-5*, *osm-6*, *bbs-5* and *xbx-1* was previously described, as referenced.

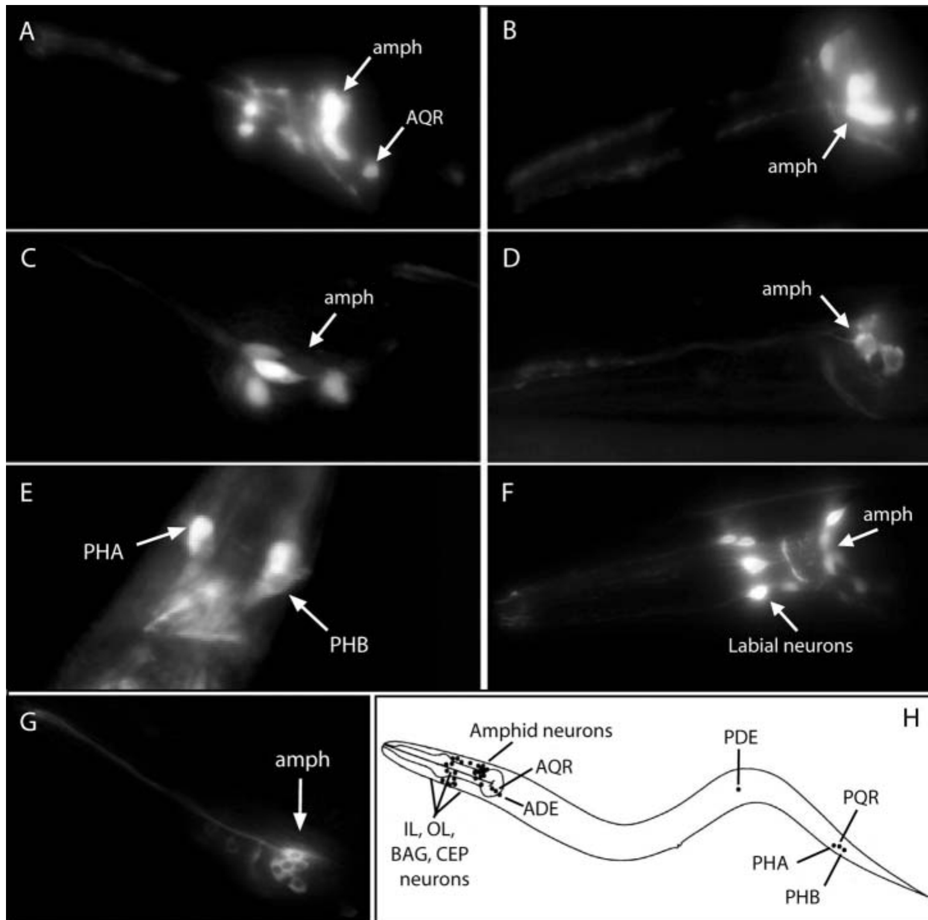


Fig. 3. Newly discovered X-box genes. (A) *xbx-2::gfp* is expressed in most or all CSN. (B) *xbx-3::gfp* is expressed in amphid and phasmid (not shown) neurons. (C) *xbx-4::gfp* is expressed in a subset of amphid neurons. (D) *xbx-5::gfp* is expressed in a subset of amphid neurons. (E) *xbx-6::gfp* is expressed in many different types of cells, including the phasmid neurons PHA and PHB. (F) *xbx-7::gfp* is expressed in a subset of amphid and interlabial neurons. (G) *tub-1::gfp* is expressed in most or all CSN. (H) Schematic diagram of CSN positioning in the *C. elegans* hermaphrodite (after Collet et al., 1998).

intraflagellar transport (IFT): *osm-1, osm-5, osm-6, che-2, che-13, xbx-1.* The mechanism of IFT was originally described in the biflagellate alga *Chlamydomonas reinhardtii* (Kozminski et al., 1993). It is characterized by the movement of IFT particles along ciliary/flagellar axonemal microtubules by means of kinesin and dynein motor molecules. *C. elegans xbx* genes encoding IFT proteins are well described. IFT components such as OSM-1, OSM-5, OSM-6, and CHE-13 are associated with the heterotrimeric motor protein kinesin-II. They are essential for anterograde transport (Signor et al., 1999; Haycraft et al., 2001; Haycraft et al., 2003). The gene *xbx-1* encodes a dynein light intermediate chain (DLIC) that is important for retrograde transport within cilia (Schafer et al., 2003).

Herein we report a new IFT gene, *xbx-2* (D1009.5). The *xbx-2::gfp* promoter fusion was strongly expressed in most of the ciliated sensory organs of the worm – amphids, phasmids, inner and outer labial quadrants (Fig. 3A). The XBX-2 protein contains a Tctex-1 domain that belongs to the family of dynein light chain (DLC) proteins. These molecules are essential for dynein assembly and participate in specific motor-cargo interactions (DiBella et al., 2001; Tai et al., 2001). To analyze the possible role of XBX-2 in the IFT process, we generated transgenic worms expressing XBX-2::GFP protein. Using time-lapse confocal microscopy, we observed movement of XBX-2::GFP particles along the ciliary axoneme in both retrograde (Fig. 4) and anterograde directions (see also Movie 1 in supplementary material).

A major part of the first group was also composed of *C. elegans* orthologs of human Bardet–Biedl syndrome genes: *bbs-1*, *bbs-2*, *bbs-5*, *bbs-7* and *bbs-8*. Recent data suggest that *bbs* genes are probably involved in the development of the basal body during cilium formation (Ansley et al., 2003; Li et al., 2004) and are required for assembly and proper function of some IFT components (Blacque et al., 2004; Fan et al., 2004). We expressed *bbs* gene members both in wild-type and in *daf-19* mutant backgrounds (in parallel with X-box mutagenesis experiments) and found that this group of genes strongly requires both DAF-19 function and proper X-box composition (Table 1). These results confirm that genes encoding general cilium structure molecules are strongly regulated by DAF-19.

The second group of *xbx* genes includes many novel genes, which are expressed in various subsets of CSN. For example, *xbx-3::gfp* was strongly expressed in all amphid and phasmid neurons, but not in other ciliated sensilla (Fig. 3B). In *daf-19* mutant worms, expression was restricted to one amphid neuron and abolished in phasmids, with occasional ectopic expression in other tissues. Expression of the novel gene *xbx-4* was also observed in some amphid and phasmid neurons (Fig. 3C), but in *daf-19* mutants it was completely abolished in both organs. The *xbx-5::gfp* promoter fusion was characterized by faint, punctate expression in phasmids and some amphid neurons (Fig. 3D). In a *daf-19* mutant background expression was absent in amphids but still visible in phasmids. The predicted XBX-5 protein contains seven transmembrane domains and can be considered as a possible receptor. The *xbx-6::gfp* construct was abundantly expressed in many cell types: pharyngeal muscles, numerous neurons in the head and tail regions, the ventral nerve cord and body wall muscles. Because of high overall expression levels, we were not able to identify individual CSN in the head region, but we observed expression in phasmids (Fig. 3E), which was strictly DAF-19 dependent. The expression in other cells was unchanged in *daf-19* mutants. The *xbx-6* gene encodes an *N*-methyl-D-aspartate receptor-associated protein. The novel gene *xbx-7* was expressed in phasmids, some amphid neurons

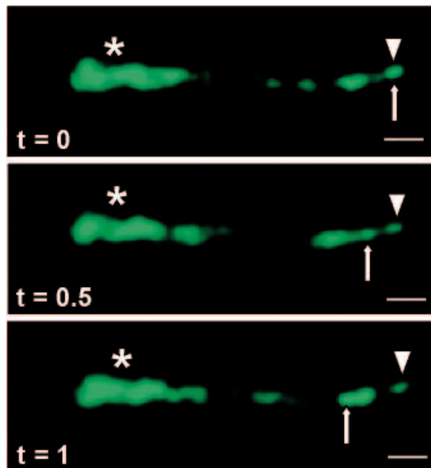


Fig. 4. An example of retrograde movement for XBX-2::GFP particles in a phasmid cilium (see also Movie 1 in supplementary material). The ciliary transition zone is marked with an asterisk. The arrowhead indicates the initial position of the moving particle at time zero (t=0). The arrow indicates the position of the moving particle at different time points (t=0.5 and 1 seconds). Scale bars: 1 μm.

and in interlabial neurons (Fig. 3F), while in *daf-19* mutants expression was significantly reduced. The gene *che-11* encodes a large protein, which is orthologous to *Chlamydomonas* IFT140 (Qin et al., 2001). The expression level of *che-11::gfp* was significantly reduced in phasmids and amphids in *daf-19* mutants (Table 2), while in the labial neurons it was almost unchanged. The nuclear hormone receptor gene *nhr-44* was expressed in head neurons, including the ciliated sensory neuron ASK, and also in other cell types, and its overall expression properties fit those of other *xbx* genes (Table 1).

Within the second group of *xbx* genes we especially note the genes *odr-4*, *nud-1* and *tub-1*, the molecular identities of which were previously described.

The gene *odr-4* has been shown to be important for the localization of some seven transmembrane domain odorant receptors to cilia (Dwyer et al., 1998). The expression of an ODR-4::GFP translational fusion was significantly reduced both in *daf-19* mutants and after mutation of the X-box sequence (data not shown). This indicates that not only transport mechanisms within cilia (IFT), but also to cilia (ODR-4) are under DAF-19 control.

The *C. elegans* ortholog of the *NudC* gene of the fungus *Aspergillus nidulans*, *nud-1*, was identified as a candidate *xbx* gene during our search. NUD-1 is an important component in microtubule-dependent nuclear positioning, which is required for proper growth, development and cellular function in both lower and higher eukaryotes. Sustained expression of *nud-1::gfp* in CSN was previously described (Dawe et al., 2001). We introduced this construct into a *daf-19* mutant background and found its expression drastically reduced (Table 2).

The gene *tub-1* is the worm ortholog of murine *tubby* which, when mutated, leads to neuronal deficits and late-onset obesity (Carroll et al., 2004). *tub-1* mutant worms exhibit functional defects in CSN and show a mild elevation of lipid accumulation (H.Y.M., unpublished data). The translational GFP fusion for *tub-1* was expressed only in the cytoplasm in all CSN (Fig. 1G). The expression level of this construct was significantly reduced in a *daf-19* mutant background (Table 2). We analyzed the 5′-UTR region of the human TUB gene and found two X-box-like sequences that perfectly match the *C. elegans* consensus: GTTGCC AT GGAAAC (–296) and GTTGCT AT AGTAAC (–339). Intriguingly, microtubule-associated protein 1A (MAP1A) can modify hearing defects in *tubby* mice (Ikeda et al., 2002), and its expression is regulated by RFX molecules (Nakayama et al., 2003). These observations suggest that the regulation of *tubby* pathways by RFX transcription factors can be conserved in evolution. Based on the data from the second group of *xbx* genes, we conclude that DAF-19 only partially regulates the expression of certain genes. These genes may only be required for specialized functions in CSN or during ciliogenesis.

The remaining genes analyzed were mostly expressed in many other different cell types and only in very few cases was expression observed predominantly in CSN (Table 1; see Table S2 in supplementary material). We checked some genes from this group (*zag-1*, *aqp-2*, *gpa-9*, F55D12.1, F17A2.3) in a *daf-19* mutant background and found that expression patterns were unchanged in the absence of DAF-19 function. Most of these X-box matches differ from the refined consensus or are located further upstream of the ATG (Table 1).

Table 2. Expression properties of novel *xbx* genes in wild type and in *daf-19(m86)* backgrounds*

	Number of amphid neurons expressing GFP						Average number of expressing amphid neurons	Number of phasmid neurons expressing GFP			Average number of expressing phasmid neurons
Genotype	0	1-5	6-10	11-15	16-20	21-24		0	1-2	3-4	
che-11::gfp											
Wild type	–	–	–	–	99	1	19	–	1	99	4
daf-19(m86)	61	39	–	–	–	–	1	87	12	1	1
nud-1::gfp											
Wild type	–	12	88	–	–	–	7	1	74	25	2
daf-19(m86)	100	–	–	–	–	–	0	100	–	–	0
tub-1::gfp											
Wild type	–	–	–	9	88	3	18	–	–	100	4
daf-19(m86)	–	–	91	9	–	–	8	19	75	6	2

*Data are given as percentage expression in different numbers of amphid and phasmid neurons. The average number of expressing neurons is specified. Two independent transgenic lines were analyzed for each gene.

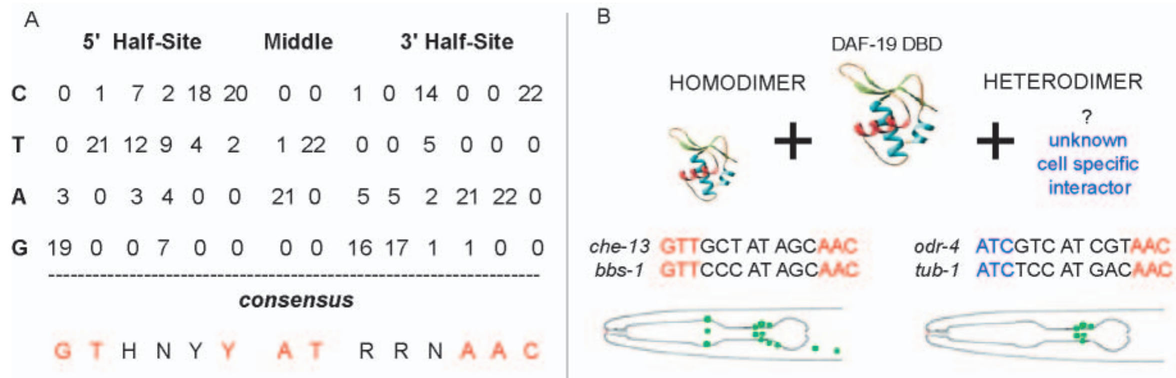


Fig. 5. Properties of the X-box motif in *C. elegans.* (A) The X-box consensus sequence obtained from in vivo expression analysis. Nucleotides marked in red are strongly conserved and important for the proper function of the motif. (B) Proposed scheme for the difference in expression patterns observed for *xbx* genes. Depending on the nucleotide composition of the X-box, DAF-19 can bind to the motif either in homodimer (driving expression in many, most or all CSN) or in heterodimer form, interacting together with some cell-specific factor (driving expression in a subset of CSN).

An additional confirmation and filtering mechanism used was a cross-species comparison, where candidates were classified as likely or unlikely *xbx* genes depending on the conservation of the X-box motif sequence between *C. elegans* and *C. briggsae* (Table 1). We found that most of the genes from the first group have putative *C. briggsae* orthologs with nearly identical X-box sequences and positions in their promoters. In the second group, the frequency of X-box occurrences in *C. briggsae* orthologs was reduced. *C. elegans xbx* gene candidates that were expressed predominantly in cell types other than CSN typically have no X-box-like sequences in promoters of their *C. briggsae* orthologs (Table 1; see Table S2 in supplementary material). Thus, the actual X-box sequence and position conservation between *C. elegans* and *C. briggsae* (and maybe other organisms) can be used as an additional measure to deduce possible molecular roles of *xbx* genes in cilia or CSN.

The expression data from 22 GFP fusions that show dependence on DAF-19 function and actual X-box sequence (Table 1) allowed us to derive an in vivo refined consensus for *C. elegans xbx* genes (Fig. 5A). This consensus can be characterized as a 14 nt imperfect palindromic sequence (GTHNYY AT RRNAAC) consisting of conserved nucleotides at 5′ (GT) and 3′ (AC) ends of the two half-sites separated by a conserved AT spacer. Most of the X-boxes matching this consensus are located in immediate proximity to the gene start (in the range of 50-250 bp upstream of the ATG). Lack of DAF-19 function or changes in the actual X-box sequence lead to drastic reduction or variation of gene expression patterns, suggesting a crucial role of the given motif sequence in the regulation of target genes (see also Discussion).

Functional analysis of *xbx* genes

Numerous mutations have been generated in *C. elegans* that affect cilia and their formation (Perkins et al., 1986; Starich et al., 1995). They cause various mutant phenotypes, including Dyf (fluorescent dye filling defective), Osm (avoidance of high osmotic strength defective), Mat (male mating defective), Che and Odr (chemo- and odorant sensation defective).

xbx genes from group 1 show Dyf and to various extents also different sensory phenotypes (Table 1) (Collet et al., 1998; Signor et al., 1999; Fujiwara et al., 1999; Haycraft et al., 2001; Haycraft et al., 2003; Schafer et al., 2003; Blacque et al., 2004; Li et al., 2004). Mutations that reduce dye filling of amphid and phasmid neurons are indicative of general defects in cilium structure and are often accompanied by various sensory mutant phenotypes (Starich et al., 1995).

The availability of a large group of *xbx* gene candidates obtained in our computational search prompted us to try a screening approach for new members of general cilium structure and function using genetic interference mediated by double stranded RNA (RNAi). We analyzed already known *dyf* genes (*che-13*, *osm-5*) together with novel candidates (*bbs-2*, *bbs-7*, *xbx-1*, *xbx-2*) with regard to the Dyf phenotype. It was known that RNAi is less efficient in neuronal types of cells (Simmer et al., 2002). To increase possible effects of interference, in parallel to wild-type worms we also tested *rrf-3* mutants, which are sensitive to RNAi in diverse tissues, especially in neurons (Simmer et al., 2002). We found that RNAi of the ciliary genes tested did not result in strong Dyf phenotypes in wild type or *rrf-3* mutants (data not shown). Therefore, RNAi cannot be implemented as a quick and easy screening technique with regard to the Dyf phenotype in *C. elegans*.

Unlike those in group 1, the roles of many genes in group 2 are largely unknown. To begin the functional investigation of the second group we analyzed two genes, *xbx-4* and *xbx-6*, mutants of which were available. The *xbx-4(ok635)* deletion extends over 951 bp starting in the promoter region and ending in the first intron, completely eliminating the beginning of the gene. The *xbx-6(ok852)* deletion extends over 1720 bp starting in the promoter region and covering five of the six exons of the gene. Since the expression of these genes is associated with CSN, we first focused our efforts on phenotypes related to defects in cilium structure or sensory abnormalities. The following results were obtained: both analyzed deletion alleles demonstrated wild-type responses with regard to fluorescent dye filling, high osmotic strength avoidance and in odorant sensation assays using three different odors (data not shown).

These *xbx-4* and *xbx-6* results suggest that in some instances members of the *xbx* gene family may have specialized molecular functions and therefore mutants have more

Development

Development

specialized sets of sensory phenotypes, although we cannot formally exclude genetic redundancy with other (*xbx*) genes expressed in CSN. For example, two other members of group 2, the genes *odr-4* and *tub-1*, when mutated, also do not produce general structural defects of cilia, but more specialized functional ciliary abnormalities, such as selective defects in odorant sensation (Dwyer et al., 1998) (H.Y.M., unpublished).

In conclusion, our data support the sorting of *xbx* genes into different groups, where members of group 1 are typically required for more general aspects of cilia formation, while genes from group 2 are typically required for more specialized functions within cilia and/or CSN.

Discussion

Efficiency of ciliary gene searching

To find groups of genes that belong to general and to specialized ciliary gene classes we used a promoter motif, the X-box, which participates in the regulation of certain ciliary genes in the nematode *Caenorhabditis elegans*. To make our search as efficient as possible we repeated, expanded and refined it with new sets of parameters (in a bootstrap-like fashion). As a result, we extracted from the *C. elegans* genome a list of 758 *xbx* gene candidates (Table S1 in supplementary material). We predict around 150 candidates from our list to be ciliary/ciliated sensory neuron (CSN)-specific genes, since they meet the following combination of criteria, found through experimental work:

(i) Out of more than 30 *xbx* genes tested for GFP expression patterns, about 60% of them, having an X-box motif fitting the refined consensus, were expressed in CSN, and most of them only in this group of neurons. When tested by mutational analyses expression in CSN was nearly always DAF-19 and X-box dependent.

(ii) Most of these experimentally confirmed *xbx* genes have the X-box promoter motif fairly close upstream of the ATG (up to –250), a pattern found previously for the first few cilium-specific *xbx* genes analyzed (Swoboda et al., 2000). By using various randomly scrambled X-box sequences for searches we found that the concentration of hits closely upstream of the ATG disappeared and became uniform throughout the search space (Fig. 1D). We also analyzed available expression data for genes, where the X-box match was located further upstream of the ATG and found that these genes are typically not expressed in a ciliary- or CSN-specific manner (Table S2 in supplementary material). By using the X-box as an anchor it is possible that gene predictions for some of our candidates could be re-evaluated. Therefore, some X-box matches located further upstream can still be considered as functional promoter motifs (Table S1 in supplementary material).

(iii) Finally, cross-species comparisons with the closely related nematode *C. briggsae* show strong conservation of the X-box sequence and position within promoters of many of the *C. elegans* experimentally confirmed *xbx* genes.

We note that our search is not completely unbiased, because the search algorithm is based on an originally small set of experimentally proven *xbx* genes. Therefore, as further *xbx* genes will be shown to be ciliary- or CSN-specific genes, their X-boxes will be included into the search parameters (bootstrap mechanism).

Other genome search approaches, in part utilizing X-box matches as a parameter, were used in different organisms to find general components of cilia formation (Li et al., 2004; Avidor-Reiss et al., 2004), yielding a set of conserved ciliary genes and gene candidates. We compared information from flagellar and basal body genes (Li et al., 2004) with our list of X-boxes and found an overlap of 15 X-box matches. At the same time, we found eight additional *xbx* gene candidates (Table S1 in supplementary material). Despite overlaps, many of the X-box matches were different between the respective searches, possibly because of different experimental parameters. Another important comparison was obtained from a recent study of olfactory neuron-specific genes (Colosimo et al., 2004), where we found 56 genes in common (Table S1 in supplementary material). All approaches together with further filtering mechanisms will give a complete, exhaustive list of genes important for structure and function of cilia and CSN.

Properties of the X-box motif in *C. elegans*

cis-Regulatory elements are information processing devices hardwired into the genomic DNA sequence, the function of which is to regulate gene expression (Davidson, 2001). Frequently, they are organized into modules that include many sites for DNA binding proteins (Howard and Davidson, 2004). In *C. elegans*, only some cell-type-defining transcription factors (CEH-10/TTX-3, MEC-3) target single binding sites in promoters of regulated genes (Zhang et al., 2002; Wenick and Hobert, 2004). Typically, also DAF-19 targets contain only a single X-box motif in their promoters. Nevertheless, different *xbx* genes show different expression properties, suggesting the presence of specific DAF-19 co-regulators. We propose that the information about particular gene expression profiles could already be stored at the level of the X-box sequence. For example, all genes for cilium structure and IFT from group 1 have perfect matches to the refined X-box motif consensus. Each gene of this group is expressed in most or all CSN and is strongly dependent on DAF-19 function. Whereas a perfect match to the 'refined consensus' does not automatically predict expression in all CSN, variations in the X-box motif sequence, especially in the more variable half-site (GTHNYY), predict different gene expression properties (*xbx-3*, *che-11*, *odr-4*, *tub-1*). These genes are either expressed only in a subset of CSN (*xbx-3*, *odr-4*) or are only partially dependent on DAF-19 function (*che-11*, *tub-1*).

Structural experiments have shown that each half-site of a symmetric X-box interacts with both DBDs of the RFX homodimer (Gajiwala et al., 2000). Nevertheless, binding of RFX is not ultimately dependent on dimerization and monomers can bind to a single 'high-affinity' half-site (RGYAAC) (Siegrist et al., 1993; Emery et al., 1996b).

We hypothesize that DAF-19 is a crucial transcription factor of genes required for general cilium formation. In this case, DAF-19 recognizes symmetric X-boxes as a homodimer and strongly activates their expression in most or all CSN (Fig. 5B). If a target gene contains a more asymmetric X-box sequence, it can be recognized by heterodimers of DAF-19 together with other, as yet unidentified factors, leading then to specific expression patterns in subsets of CSN (Fig. 5B). Hypothetically, X-box motif distances from the ATG could also contribute to target gene expression variability.

DAF-19 regulates the development of a 'ciliary module'

The development of sensory neurons in *C. elegans* is a complicated process that includes many stages and interactions of different transcription factors (Melkman and Sengupta, 2004). The gene *daf-19* acts at late stages of sensory neuron development when the respective cell fates have already been determined and specification and subsequent differentiation occurs. Several different types of genes are required to produce functional cilia in the worm: genes for their molecular structure, genes implicated in ciliary transport machineries and genes involved in signal reception and transduction (Jansen et al., 1999; Troemel, 1999; Rosenbaum and Witman, 2002; Melkman and Sengupta, 2004).

Our previous model associated DAF-19 regulation with only a certain group of genes functioning in cilium morphogenesis and architecture (Swoboda et al., 2000). The data obtained in our current research suggest that the repertoire of DAF-19-dependent genes is much broader (Fig. 2). For the first time we have found that DAF-19 can also regulate genes of ciliary function. For example, expression of the gene *odr-4* requires both a correct X-box sequence and the presence of DAF-19. The ODR-4 protein has been shown to be an important factor for localizing a subset of seven transmembrane domain odorant receptors to cilia (Dwyer et al., 1998). Moreover, we have shown that DAF-19 can directly regulate expression levels of some putative receptor proteins. For example, the genes *xbx-5* and *xbx-6* encode a seven transmembrane domain protein and an *N*-methyl-D-aspartate receptor-associated protein, respectively.

In addition to the group of genes with signal reception properties, we found X-boxes in promoters of proposed transcription factors (Fig. 2). Most of these factors contain a C2H2-type zinc-finger domain. The important role of this type of transcription factors for the development of cell-specific properties in CSN was already described for the gene *che-1* (Uchida et al., 2003). Thus, we predict the presence of cilium-specific developmental cascades directed by DAF-19-dependent transcription factors. These cascades may be required for specialized ciliary functions as well as being necessary for the possible parallel regulation of CSN specification and their final functional differentiation. For example, it has been demonstrated that DAF-19 can affect the expression of indirect targets in the HOB-specific pathway through some unknown factor(s) (the male-specific ciliated HOB neuron is necessary for sensation of the hermaphrodite vulva during mating) (Yu et al., 2003). Our own data also suggest that DAF-19 could be required not only for general cilia formation, but in some instances for the development of cell-specific properties as well. For example, we observed that the gene *nhr-44* was expressed in a DAF-19-dependent manner in the ciliated sensory neuron ASK. This gene belongs to the nuclear hormone receptor family, which includes many ligand-regulated transcriptional modulators involved in many developmental processes (Miyabayashi et al., 1999).

Another example is the gene *nud-1*. We propose that certain microtubule-associated molecules (like NUD-1) acting during early developmental stages could later be recruited by DAF-19 for the purposes of cilia development. The important role of *nud-1* in nuclear migration during embryogenesis in *C. elegans* was previously described (Dawe et al., 2001). It has also been shown that the mammalian ortholog of NUD-1, NudC, associates with the dynein motor complex during neuronal migration (Aumais et al., 2001). Therefore, we suggest a possible role for NUD-1 as a component of IFT during ciliogenesis. In addition to *nud-1*, we extracted two further X-box-containing genes from our list of candidates, *spd-5* (Hamill et al., 2002) and *dlc-1*, which are also involved in nuclear migration during early embryonic development and might later be recruited by DAF-19 for the development of CSN.

Considering the above we propose a model where DAF-19 regulates the development of a 'ciliary module' during the differentiation of sensory neurons in *C. elegans* (Fig. 6). According to this model, DAF-19 is a key factor for the general development of cilia. At the same time, together with other factors, it can drive the expression of genes required for specialized functions in cilia.

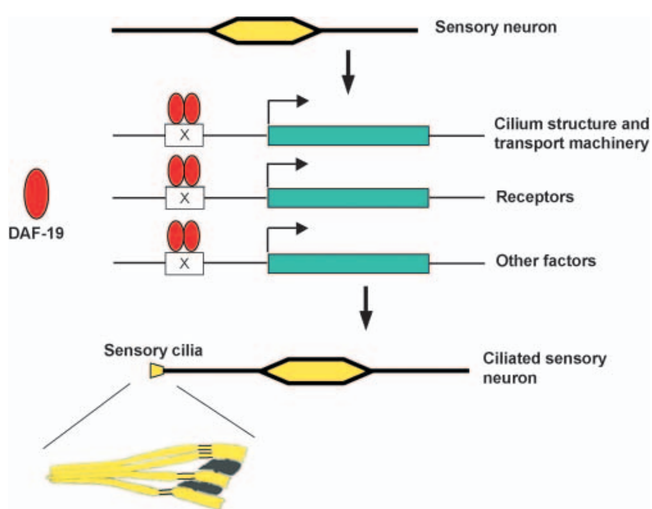


Fig. 6. Development of a 'ciliary module' in *C. elegans*. DAF-19 regulates the development of the module, which includes genes for the cilium structure and transport machinery, receptors and other factors. The activation of this module leads to the formation of functional ciliated endings during specification of sensory neurons in the worm.

xbx genes and cilia-dependent diseases

In mammals cilia are near ubiquitous organelles that project from the surfaces of many different cell types to carry out motility and sensory functions (Afzelius, 2004). Cilia have been implicated directly in many developmental processes such as generation of left-right asymmetry, heart development, maintenance of the renal epithelium, respiratory function, and physiological balance of the cerebrospinal fluid. Defects in cilia function and structure lead to a wide range of developmental problems and diseases (Pazour and Rosenbaum, 2002; Afzelius, 2004), among them: Bardet–Biedl syndrome, polycystic kidney disease (PKD), X-linked retinitis pigmentosa 2, nephronophthisis, maturity-onset obesity, etc. Many *C. elegans* genes that are orthologous to genes of human cilia-dependent diseases are X-box-containing genes. Therefore, we believe that the understanding of ciliogenesis in *C. elegans* will have a significant impact on the understanding and treatment of cilia-based pathologies in humans.

Development

We thank the following people for technical assistance, helpful discussions, unpublished information, *C. elegans* strains, cosmid clones, cDNA clones, RNAi clones and GFP vectors: J. Ahringer, M. Barr, G. Caldwell, A. Coulson, A. Fire, C. Haycraft, Y. Kohara, G. Merrihew, E. Onischenko, S. Tat, H. Tian, B. Yoder and especially members of our and the T. Bürglin laboratories. We thank G. Gahmon and A. Laurençon for critical reading of the manuscript. We thank the C. elegans Genome Sequencing Consortium for providing genome sequence information and the Caenorhabditis Genetics Center, which is funded by the National Institutes of Health, for providing some of the *C. elegans* strains used in this study. We thank the C. elegans Knockout Consortium for providing gene deletion mutants. Work by H.Y.M. was supported by a postdoctoral fellowship from the Human Frontier Science Promotion Organization (Strasbourg, France). M.R.L. acknowledges scholar awards from the Canadian Institutes of Health Research (CIHR) and Michael Smith Foundation for Health Research (MSFHR), and funding provided by the Heart and Stroke Foundation of B. C. and Yukon and the Natural Sciences and Engineering Research Council of Canada (NSERC). This work was supported in part by Public Health Service Grants from NIH to J.H.T. and G.R., and by a grant to P.S. from the Swedish Foundation for Strategic Research (SSF) in Stockholm, Sweden.

Supplementary material

Supplementary material for this article is available at http://dev.biologists.org/cgi/content/full/132/8/1923/DC1

Note added in proof

The *C. elegans* gene names *xbx-2* and *dylt-2* describe the same gene, *D1009.5*.

References

Afzelius, B. A. (2004). Cilia-related diseases. *J. Pathol.* **204**, 470-477.

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., Mah, A. K., Johnsen, R. C., Cavender, J. C., Lewis, R. A., Leroux, M. R., Beales, P. L. and Katsanis, N. (2003). Basal body dysfunction is a likely cause of pleiotropic Bardet–Biedl syndrome. *Nature* **425**, 628-633.

Aumais, J. P., Tunstead, J. R., McNeil, R. S., Schaar, B. T., McConnell, S. K., Lin, S., Clark, G. D. and Yu-Lee, L. (2001). NudC associates with Lis1 and the dynein motor at the leading pole of neurons. *J. Neurosci.* **21**, 1-7.

Avidor-Reiss, T., Maer, A. M., Koundakjian, E., Polyanovsky, A., Keil, T., Subramaniam, S. and Zuker, C. S. (2004). Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. *Cell* **117**, 527–539.

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., Johnsen, R. C., Audeh, M., Plasterk, R. H. A., Baillie, D. L., Katsanis, N., Quarmby, L. M., Wicks, S. R. and Leroux, M. R. (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.

Bonnafe, E., Touka, M., AitLounis, A., Baas, D., Barras, E., Ucla, C., Moreau, A., Flamant, F., Dubruille, R., Couble, P., Collignon, J., Durand, B. and Reith, W. (2004). The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. *Mol. Cell. Biol.* **10**, 4417-4427.

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

Carroll, K., Gomez, C. and Shapiro, L. (2004). Tubby proteins: the plot thickens. *Mol. Cell. Biol.* **5**, 55-63.

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.

Colosimo, M. E., Brown, A., Mukhopadhyay, S., Gabel, C., Lanjuin, A. E., Samuel, A. D. and Sengupta, P. (2004). Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types. *Curr. Biol.* **14**, 2245-2251.

DiBella, L. M., Benashski, S. E., Tedford, H. W., Harrison, A., Patel-King, R. S. and King, S. M. (2001). The Tctex1/Tctex2 class of dynein light chains. *J. Biol. Chem.* **276**, 14366-14373.

Davidson, E. H. (2001). *Genomic Regulatory Systems*. San Diego: Academic Press.

Dawe, A. L., Caldwell, K. A., Harris, P. M., Morris, N. R. and Caldwell, G. A. (2001). Evolutionarily conserved nuclear migration genes required for early embryonic development in *Caenorhabditis elegans. Dev. Genes. Evol.* **211**, 434-441.

Dubruille, R., Laurencon, A., Vandaele, C., Shishido, E., Coulon-Bublex, M., Swoboda, P., Couble, P., Kernan, M. and Durand, B. (2002). Drosophila Regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* **129**, 5487-5498.

Dutcher, S. K. (1995). Flagellar assembly in two hundred and fifty easy-to-follow steps. *Trends Genet.* **11**, 398-404.

Dwyer, N. D., Troemel, E. R., Sengupta, P. and Bargmann, C. I. (1998). Odorant receptor localization to olfactory cilia is mediated by ODR-4, a novel membrane-associated protein. *Cell* **93**, 455-466.

Emery, P., Durand, B., Mach, B. and Reith, W. (1996a). RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucl. Acids Res.* **24**, 803-807.

Emery, P., Strubin, M., Hofmann, K., Bucher, P., Mach, B. and Reith, W. (1996b). A Consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity. *Mol. Cell. Biol.* **16**, 4486-4494.

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., Green, J. S., Lewis, R. A., Haelst, M. M., Parfrey, P. S., Baillie, D. L., Beales, P. L., Katsanis, N., Davidson, W. S. and Leroux, M. R. (2004). Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet-Biedl syndrome. *Nat. Genet.* **36**, 989-993.

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.

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.

Hamill, D. R., Severson, A. F., Carter, J. C. and Bowerman, B. (2002). Centrosome maturation and mitotic spindle assembly in *C. elegans* require SPD-5, a protein with multiple coiled-coil domains. *Dev. Cell* **3**, 673-684.

Han, Y., Kwok, B. H. and Kernan, M. J. (2003). Intraflagellar transport is required in *Drosophila* to differentiate sensory cilia but not sperm. *Curr. Biol.* **13**, 1679-1686.

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.

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.

Howard, M. L. and Davidson, E. H. (2004). cis-Regulatory control circuits in development. *Dev. Biol.* **271**, 109-118.

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.

Ikeda, A., Zheng, Q. Y., Zuberi, A. R., Johnson, K. R., Naggert, J. K. and Nishina, P. M. (2002). Microtubule-associated protein 1A is a modifier of tubby hearing (*moth1*). *Nat. Genet.* **30**, 401-405.

Jansen, G., Thijssen, K. L., Werner, P., Horst, M., Hazendonk, E. and Plasterk, R. H. A. (1999). The complete family of genes encoding G proteins of *Caenorhabditis elegans. Nat. Genet.* **21**, 414-419.

Kozminski, K. G., Johnson, K. A., Forscher, P. and Rosenbaum, J. L. (1993). A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proc. Natl. Acad. Sci. USA* **90**, 5519-5523.

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., Lewis, R. A., Green, J. S., Parfrey, P. S., Leroux, M. R., Davidson, W. S., Beales, P. L., Guay-Woodford, L. M., Yoder, B. K., Stormo, G. D., Katsanis, N. and Dutcher, S. K. (2004). Comparative genomics identifies a flagellar and basal body proteome that includes the *BBS5* human disease gene. *Cell* **117**, 541-552.

Maijgren, S., Sur, I., Nilsson, M. and Toftgard, R. (2004). Involvement of RFX proteins in transcriptional activation from a Ras-responsive enhancer element. *Arch. Dermatol. Res.* **295**, 482-489.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991).

Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959-3970.

Melkman, T. and Sengupta, P. (2004). The worm's sense of smell. Development of functional diversity in the chemosensory system of *Caenorhabditis elegans*. *Dev. Biol.* **265**, 302-319.

Miyabayashi, T., Palfreyman, M. T., Sluder, A. I., Slack, F. and Sengupta, P. (1999). Expression and function of members of a divergent nuclear receptor family in *Caenorhabditis elegans*. *Dev. Biol.* **215**, 314-331.

Nakayama, A., Murakami, H., Maeyama, N., Yamashiro, N., Sakakibara, A. 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.

Otsuki, K., Hayashi, Y., Kato, M., Yoshida, H. and Yamaguchi, M. (2004). Characterization of dRFX2, a novel RFX family protein in Drosophila. *Nucl. Acids Res.* **32**, 5636-5648.

Pazour, C. J. and Rosenbaum, J. L. (2002). Intraflagellar transport and cilia-dependent diseases. *Trends Cell Biol.* **12**, 551-555.

Perkins, L. S., Hedgecock, E. M., Thomson, J. N. and Culotti, J. G. (1986). Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Dev. Biol*. **117**, 456-487.

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.

Reith, W. and Mach, B. (2001). The Bare lymphocyte syndrome and the regulation of MHC expression. *Annu. Rev. Immunol.* **19**, 331-373.

Rosenbaum, J. L. and Witman, G. B. (2002). Intraflagellar transport. *Nature Reviews: Mol. Cell. Biol.* **3**, 813-825.

Ruvinsky, I. and Ruvkun, G. (2003). Functional tests of enhancer conservation between distantly related species. *Development* **130**, 5133-5142.

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.

Siegrist, C. A., Durand, B., Emery, P., David, E., Hearing, P., Mach, B. and Reith, W. (1993). RFX1 is identical to EF-C and functions as a transactivator of the hepatitis B virus enhancer. *Mol. Cell. Biol.* **13**, 6375-6384.

Signor, D., Wedaman, K. P., Orozco, J. T., Dwyer, N. D., Bargmann, C. I., Rose, L. S. and Scholey, J. M. (1999). Role of a class DHC1b dynein in retrograde transport of IFT motors and IFT raft particles along cilia, but not dendrites, in chemosensory neurons of living *Caenorhabditis elegans*. *J. Cell. Biol.* **147**, 519-530.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J. and Plasterk, R. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Curr. Biol.* **12**, 1317-1319.

Starich, T. A., Herman, K. R., Kari, C. K., Yeh, W., 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.

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**, 166-192.

Swoboda, P., Adler, H. 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.

Tai, A. W., Chuang, J. and Sung, C. (2001). Cytoplasmic dynein regulation by subunit heterogeneity and its role in apical transport. *J. Cell. Biol.* **153**, 1499-1509.

Timmons, L., Court, D. L. and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263**, 103-112.

Troemel, E. R. (1999). Chemosensory signaling in *C. elegans*. *BioEssays* **21**, 1011-1020.

Uchida, O., Nakano, H., Koga, M. and Ohshima, Y. (2003). The *C. elegans che-1* gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons. *Development* **130**, 1215-1224.

Wacker, I., Schwarz, V., Hedgecock, E. M. and Hutter, H. (2003). *zag-1*, a Zn-finger homeodomain transcription factor controlling neuronal differentiation and axon outgrowth in *C. elegans*. *Development* **130**, 3795-3805.

Ward, S., Thomson, N., White, J. G. and Brenner, S. (1975). Electron microscopical reconstruction of the anterior sensory anatomy of the nematode *Caenorhabditis elegans*. *J. Comp. Neurol*. **160**, 313-337.

Wenick, A. S. and Hobert, O. (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., Southgate, E., Thomson, J. N. and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **314**, 1-340.

Wu, S. Y. and McLeod, M. (1995). The sak1 gene of *Schizosaccharomyces pombe* encodes an RFX family DNA-binding protein that positively regulates cyclic AMP-dependent protein kinase-mediated exit from the mitotic cell cycle. *Mol. Cell. Biol.* **15**, 1479-1488.

Yu, H., Pretot, R. F., Bürglin, T. R. and Sternberg, P. W. (2003). Distinct roles of transcription factors EGL-46 and DAF-19 in specifying the functionality of a polycystin-expressing sensory neuron necessary for *C. elegans* male vulva location behavior. *Development* **130**, 5217-5227.

Zhang, Y., Ma, C., Delohery, T., Nasipak, B., Foat, B. C., Bounoutas, A., Bussemaker, H. J., Kim, S. K. and Chalfie, M. (2002). Identification of genes expressed in *C. elegans* touch receptor neurons. *Nature* **418**, 331-335.