Forkhead transcription factor FKH-8 cooperates with RFX in the direct regulation of sensory cilia in *C. elegans*

Rebeca Brocal-Ruiz[1], Ainara Esteve-Serrano[1], Carlos Mora-Martinez[1], Maria Luisa Franco-Rivadeneira[2], Peter Swoboda[3], Juan Tena[4], Marçal Vilar[2] and Nuria Flames[1,*]

1: Developmental Neurobiology Unit, Instituto de Biomedicina de Valencia IBV-CSIC, Valencia, 46010, Spain

2: Molecular Basis of Neurodegeneration Unit, Instituto de Biomedicina de Valencia IBV-CSIC, Valencia, 46010, Spain

3: Department of Biosciences and Nutrition. Karolinska Institute. Campus Flemingsberg - NEO Building Hälsovägen 9, SE-141 83 Huddinge, Sweden

4: Centro Andaluz de Biología del Desarrollo (CABD), Consejo Superior de Investigaciones Científicas/Universidad Pablo de Olavide, Seville, 41013, Spain.

* Correspondence: nflames@ibv.csic.es

SUMMARY

Cilia, either motile or non-motile (a.k.a primary or sensory), are complex evolutionarily conserved eukaryotic structures composed of hundreds of proteins required for their assembly, structure and function that are collectively known as the ciliome. Ciliome gene mutations underlie a group of pleiotropic genetic diseases known as ciliopathies. Proper cilium function requires the tight coregulation of ciliome gene transcription, which is only fragmentarily understood. RFX transcription factors (TF) have an evolutionarily conserved role in the direct activation of ciliome genes both in motile and non-motile cilia cell-types. In vertebrates, FoxJ1 and FoxN4 Forkhead (FKH) TFs work with RFX in the direct activation of ciliome genes, exclusively in motile cilia cell-types. No additional TFs have been described to act together with RFX in primary cilia cell-types in any organism. Here we describe FKH-8, a FKH TF, as a direct regulator of the sensory ciliome genes in *Caenorhabditis elegans*. FKH-8 is expressed in all ciliated neurons in C. *elegans*, binds the regulatory regions of ciliome genes, regulates ciliome gene expression, cilium morphology and a wide range of behaviours mediated by sensory ciliated neurons. FKH-8 and DAF-19 (*C. elegans* RFX) physically interact and synergistically regulate ciliome gene expression. *C. elegans* FKH-8 function can be replaced by mouse FOXJ1 and FOXN4 but not by other members of other mouse FKH subfamilies. In conclusion, RFX and FKH TF families act jointly as direct regulators of ciliome genes also in sensory ciliated cell types suggesting that this regulatory logic could be an ancient trait predating functional cilia sub-specialization.

Keywords

cilium, transcriptional regulation, FKH, RFX, terminal differentiation, *C. elegans*

INTRODUCTION

Eukaryotic cilia are complex and highly organized organelles defined as specialized membrane protrusions formed from a stereotyped assembly of microtubules. Cilia are composed of hundreds of proteins, required for their assembly, structure and function, which are collectively known as the ciliome (**Figure 1A**). Cilia can be classified into motile or non-motile based on their function and structure: motile cilia are responsible for propelling cells or generating fluid flow while non-motile (a.k.a primary or sensory) cilia function as cellular antennae to sense extracellular stimuli (Choksi et al., 2014). Cilia appeared early in eukaryotic evolution and it is thought that in ancient unicellular eukaryotes cilia displayed mixed motile and sensory functions (Mitchell, 2017). In multicellular invertebrates, primary and motor cilia are restricted to specific cell types. In contrast, in vertebrates, primary cilia are present almost in every cell, including neurons, while motile cilia are present only in specialized cell types.
Most ciliome components are shared between motile and primary cilia and are referred as "core" ciliome (**Figure 1A**). In addition, motile cilia usually contain specialised axonemal dyneins, other motile-specific components and specific signalling proteins while the membrane of sensory cilia is decorated with receptors that trigger downstream signalling cascades when they are activated by small molecules, mechanical perturbations, or radiation.

The importance and wide range of cilia functions are underscored by the large number of congenital disorders caused by mutations in genes coding for ciliome components, which are collectively called ciliopathies (Andreu-Cervera et al., 2021; Horani and Ferkol, 2021; Lucas et al., 2020; Tobin and Beales, 2009).

These disorders cause a broad spectrum of symptoms including retinal degeneration, polycystic kidney, deafness, polydactyly, brain and skeletal malformations, infertility, morbid obesity and mental retardation. Importantly, there are still many "orphan ciliopathies", which correspond to congenital disorders classified as ciliopathies by phenotype but with yet unidentified causal mutations. Genetic variants lying in coding genes (including mutations in the ciliome genes) are easier to identify as causal mutations, however, most variants associated to human diseases lie in the non-coding genome (Chatterjee and Ahituv, 2017; GTEx Consortium et al., 2017; Timpson et al., 2018). It is currently thought that some of these non-coding variants act as regulatory mutations affecting gene expression. Thus, regulatory mutations affecting ciliome gene expression might underlie many orphan ciliopathies. Understanding the molecular mechanisms that ensure correct co-regulation of ciliome genes is then of utmost importance.

Little is known about the direct transcriptional co-regulation of ciliome gene expression (Choksi et al., 2014; Lewis and Stracker, 2020; Thomas et al., 2010). In 2000, pioneering work in *Caenorhabditis elegans* identified DAF-19, an RFX family transcription factor (TF), as a direct regulator of ciliome gene expression in the ciliated sensory neurons (Swoboda et al., 2000). This work was followed by numerous reports on the role of different members of the RFX TF family as direct ciliome gene regulators both in primary and motile cilia cell-types in several animal models including *Drosophila melanogaster*, *Danio rerio*, *Xenopus laevis* and *Mus musculus* (Ashique et al., 2009; Bonnafe et al., 2004; Chung et al., 2012; Dubruille et al., 2002; Liu et al., 2007) and in humans (Sugiaman-Trapman et al., 2018). FOXJ1, an ancient member of the Forkhead family, also acts as a direct

activator of ciliome gene transcription in several vertebrates, but its role is limited to cell types containing motile cilia (Brody et al., 2000; Chen et al., 1998; Stubbs et al., 2008; Vij et al., 2012; Yu et al., 2008). Thus, currently additional TFs acting together with RFX TFs in the direct regulation of the ciliome gene expression in sensory cilia cell-types are unknown in any organism.

Here, we take advantage of the amenability of *C. elegans* for genetic studies to understand the transcriptional regulatory logic of the non-motile primary cilliome genes. *C. elegans* contains sensory but not motile cilia. In hermaphrodites, sensory cilia are found in 25 out of the 118 neuronal types known as the ciliated sensory system (Scholey, 2007) (**Figure 1B**). We find that FKH-8, a FKH TF, is expressed in all ciliated sensory neurons in *C. elegans*, with an onset of expression concomitant to the start of ciliome gene expression. Chromatin immuno-precipitation and sequencing (ChIP-seq) data analysis shows that FKH-8 binds to a broad range of ciliome genes, at locations often near X-box motifs, the binding sites for DAF-19/RFX. *fkh-8* mutants show decreased ciliome reporter gene expression, cilia morphology abnormalities and deficits in a wide range of behaviours mediated by sensory cilia. In addition, we find FKH-8/FKH and DAF-19/RFX can physically interact and act synergistically in the regulation of ciliome genes. Finally, we show that mouse FoxJ1 and FoxN4, two ancient FKH TFs known to directly regulate ciliome gene expression in vertebrate motile-cilia cell types, rescue *fkh-8* mutant expression defects in *C. elegans*. This functional conservation is not observed with members of other FKH sub-families. Our results identify FKH-8 as the first TF acting together with RFX TFs in the direct regulation of the ciliome genes in sensory-ciliated cells and suggest that this function could be evolutionary conserved in vertebrates. Taken together, a global

ciliome regulatory logic starts to emerge in which RFX and FKH TFs could act together in the direct regulation of ciliome gene expression both in cell types containing motile or primary cilia. Considering that ancestral eukaryotic cilium is proposed to combine motile and sensory functions, we speculate that RFX / FKH regulatory module might represent the ancestral state of eukaryotic ciliome gene regulation.

RESULTS

Persistent activity of regulatory regions for ciliome genes in *daf-19*/RFX mutants

The activity of regulatory regions controlling ciliome gene expression is dramatically reduced in *daf-19(m86)* null mutants. However, for several ciliome reporters, some residual activity has been anecdotally reported (Burghoorn et al., 2012; Chu et al., 2012; De Stasio et al., 2018; Efimenko et al., 2005; Haycraft et al., 2001; Swoboda et al., 2000). As *daf-19* is the only RFX TF encoded in the *C. elegans* genome, we reasoned that persistent ciliome enhancer activity in *daf-19(m86)* null mutants would underscore the presence of additional TF families acting in concert with DAF-19. Based on previous data, we selected enhancers and built fluorescent reporters for ten phylogenetically conserved and broadly expressed core cilia components: five intraflagelar transport (IFT) genes (*che-11, osm-1, ift-20, che-13, osm-5*); the transition zone transmembrane genes *tmem-107* and *mks-1*; a Tubby family of bipartite transcription factors involved in receptor trafficking (*tub-1)*; the dynein-component *xbx-1* and the ubiquitin protein ligase *peli-1* (**Figure 1A**). Human orthologs for several of these genes are linked to ciliopathies (Horani and Ferkol, 2021; Mukhopadhyay et al., 2005; Thevenon

et al., 2016). All fluorescent reporters contain at least one experimentally validated X-box, the binding site for DAF-19/RFX (**Figure 1 Figure Supplement 1**). To avoid the dauer constitutive phenotype of *daf-19(m86)* null animals, we analysed reporter expression in *daf-19(m86); daf-12(sa204)* double mutants and *daf-12(sa204)* was used as control, as previously reported (De Stasio et al., 2018; Phirke et al., 2011).

In *daf-12(sa204)* worms, all reporters show broad activity in the ciliated system, with mean reporter expression in at least 30 ciliated neurons, except for *mks-1* and *osm-5* reporters that showed expression in less than 20 cells, suggesting other enhancers outside the analysed sequences might drive expression in additional ciliated neurons (**Figure 1C, Figure 1 Figure Supplement 2** and **Figure 1 Source data 1**). As expected, *daf-19(m86); daf-12(sa204)* double mutants show a dramatic decrease in the number of neurons positive for each reporter (**Figure 1C** and **Figure 1 Figure Supplement 2**). Importantly, all reporters except *tmem-107, mks-1* and *osm-5*, which correspond to the shortest constructs, show persistent expression in some neurons (**Figure 1C**, **Figure 1 Figure Supplement 2** and **Figure 1 Source data 1**). We hypothesised that these short constructs might lack binding sites for additional TFs acting with DAF-19. Indeed, we find that shorter versions of *xbx-1* and *peli-1* reporter constructs are more affected by *daf-19* mutation than corresponding longer constructs, consistent with shorter sequences lacking additional regulatory information (**Figure 1C**). Altogether our data strongly suggests that additional TF or TFs act together with DAF-19 to directly activate core ciliome gene expression.

Identification of FKH-8 as candidate regulator of ciliome gene expression

We reasoned that similar to *daf-19,* additional regulators of cilia gene expression could act broadly on many genes coding for ciliome components and in many different ciliated neuron types. Thus, to identify these putative candidates, we combined three strategies: *cis*-regulatory analysis of the ciliome genes, TF expression enrichment in the sensory ciliated system and TF binding to putative regulatory regions of the ciliome genes.

We built a manually curated list of 163 cilium effector genes (Materials and methods and **Figure 1 Figure Supplement 2**). This list can be divided in four categories: 1) 73 "*core components*" present in all types of cilia and thus expressed by all ciliated neurons in *C. elegans*. Core components include IFT particles, kinesins, dyneins, BBSome complex, etc; 2) 68 "*Subtype specific*" genes, that code for channels or receptors located in cilia that are expressed in a neuron type specific manner, providing neuron type specific functions; 3) 13 "*Broad expression*" genes, specifically expressed within the ciliated system but not associated with well-defined core cilia functions and 4) 9 "*Male*" genes that code for genes with male-specific cilia functions (**Figure 1 Source data 2**).

D*e novo* motif enrichment analysis using the promoters of ciliome genes identified previously known RFX consensus binding sites (X-box motif). In agreement with published results, X-box motifs are preferentially associated to "*Core"* and "*Broadly expressed"* ciliome genes (**Figure 1 Figure Supplement 3**) (Burghoorn et al., 2012; Efimenko et al., 2005; Swoboda et al., 2000). An additional motif matching the pro-neural bHLH TF *lin-32*/Atoh1 is present in 28% of the genes, with no particular bias to any ciliome category (**Figure 1 Figure Supplement 3**). The pro-neural binding motif might reflect the neuronal nature of this gene set, as in *C. elegans* cilia are only present in neurons. Motif enrichment

197 analysis limited only to the 102 genes containing predicted X-box sites identified
198 two additional motifs one showing similarity to DAF-19/RFX binding site and the
199 other to the previously reported C-BOX motif for whom TF binding has not been
200 determined (Burghoorn et al., 2012) (**Figure 1 Figure Supplement 3**). Analysis
201 limited to the 61 genes lacking predicted X-box sites failed to identify enriched
202 motifs for known transcription factors. Thus, motif enrichment of ciliome genes
203 regulatory regions failed to pinpoint additional TF candidates. The X-box motif
204 consists of a long imperfect palindromic consensus binding site with high
205 information content (**Figure 1 Figure Supplement 3**), in contrast, most TF
206 binding motifs (TFBM) are often short and degenerate, thus predicted matches
207 can be found at high frequency in the genome. This feature might underlie the
208 failure to find enriched motifs for additional TFs in ciliome gene regulatory
209 regions.
210 As an alternative to motif enrichment analysis, we turned to TF expression
211 enrichment. We hypothesized that TFs acting broadly on sensory cilia gene
212 expression could show enriched expression in the sensory ciliated neurons.
213 Using available single cell RNA expression data (sc-RNAseq) from the second
214 larval stage (Cao et al., 2017), we retrieved the expression pattern of 861 *C.*
215 *elegans* transcription factors (Narasimhan et al., 2015). Ten transcription factors
216 are specifically enriched within the ciliated sensory neurons compared to other
217 neuron types or non-neuronal tissues (**Figure 1D**). Not surprisingly, *daf-19*
218 expression is not enriched in ciliated neurons, as *daf-19* is expressed
219 panneuronally and only a specific splicing isoform is restricted to ciliated neurons.
220 The expression of TFs controlling terminal identity is often maintained throughout
221 the life of the animal. Using an independent set of sc-RNAseq data from young

adult (Taylor et al., 2021), we found that among the 10 TF candidates only FKH-8 expression is detected in all 25 different types of sensory ciliated neurons (**Figure 1 Figure Supplement 3**), suggesting it could be a good candidate to act together with DAF-19.

Finally, we interrogated 446 published ChIP-seq datasets (Luo et al., 2020), corresponding to 259 different TFs (including FKH-8 but not DAF-19), for nearby binding to the ciliome gene list (**Figure 1 Source data 2**). We find FKH-8 behaves very differently from the rest of TFs with at least one FKH-8 binding peak associated to 49% of the genes on the ciliome gene list (**Figure 1E**). FKH-8 binds both core components and subtype specific ciliome genes (**Figure 1 Figure Supplement 3**), although, similar to X-box motifs, FKH-8 binding is significantly more prevalent for core ciliome genes (75% compared to 22% binding to subtype genes). Thus, both sc-RNAseq and ChIPseq data analysis pinpoint FKH-8 as a good candidate TF to directly control ciliome gene expression.

FKH-8 is expressed in all ciliated sensory neurons

FKH-8::GFP fosmid expression at young adult stage is detected in all ciliated sensory neurons, as assessed by co-localization with the *ift-20* core ciliome reporter (**Figure 2A, B and Figure 2 Source data 1**). Three non-ciliated neurons, PVD, VC4 and VC5 show FKH-8 expression, while no expression is detected in non-neuronal tissues. *C. elegans* male nervous system contains additional ciliated sensory neurons, mostly in the tail, which also express FKH-8 (**Figure 2B**). During embryonic development, there is a similar overlap between FKH-8 and *ift-20* reporters (**Figure 2 Figure Supplement 1**). Correlation between *fkh-8*, *ift-20* and *daf-19* expression during development is also observed using

Uniform Manifold Approximation and Projection (UMAP) representation of embryonic sc-RNA-seq data (Packer et al., 2019) (**Figure 2 Figure Supplement 1**). In addition, there is a high gene expression correlation for the 73 core ciliome genes and *daf-19* or *fkh-8* expression but not with other TFs (**Figure 2 Figure Supplement 1**). Thus, our analysis shows that FKH-8 is expressed almost exclusively in the whole ciliated sensory system and its developmental expression correlates with core ciliome gene expression.

FKH-8 binds near X-boxes associated to ciliome genes

Next, we extended FKH-8 ChIPseq data analysis to the whole genome. FKH-8 binds a total of 5,035 genomic regions assigned to 3,987 genes. Most peaks are associated to promoter regions (58,65%). Gene ontology analysis of FKH-8 bound genes shows enrichment for cilia functions or dauer regulation (which is also dependent on cilia integrity) (**Figure 2C, Figure 2 Source data 2**).

DNA consensus motifs bound by FKH-8 have not been experimentally determined. FKH TF family binds the canonical consensus RYMAAYA (Pierrou et al., 1994) and an alternative motif, termed FKH-like (FHL), characterized by a GACGC core sequence (Nakagawa et al., 2013). *De novo* motif enrichment analysis of FKH-8 ChIP-seq peaks does not show any match for FKH canonical binding site but identifies a motif that highly resembles the FHL motif (**Figure 2D**). This motif, present in 27% of the peaks, is enriched at central positions suggesting it could act as FKH-8 primary binding motif (**Figure 2D**).

We noticed that eight out of the twelve functional X-boxes present in the core ciliome gene reporters analysed in Figure 1C overlap with FKH-8 ChIP-seq peaks (**Figure 1 Figure Supplement 1**). Thus, we next looked for X-box enrichment in

FKH-8 bound regions. 21% of FKH-8 peaks contain at least one match for the DAF-19 position weight matrix (**Figure 2E**). Importantly, X-boxes are preferentially found also at central locations, suggesting they could be in close proximity to FKH-8 bound sites (**Figure 2E**). X-boxes are less significantly or not significantly enriched in ChIP-seq datasets for other FKH TFs (**Figure 2 Figure Supplement 2** and **Figure 2 Source data 2**), which is consistent with specific co-binding of DAF-19 and FKH-8.

Next, we analysed the presence of X-boxes and FKH-8 binding events specifically associated to ciliome genes. Regulatory sequences for 34% of ciliome genes contain both X-box motifs and FKH-8 genomic binding (**Figure 2F** and **Figure 2 Source data 2**). This dual regulatory signature is more prevalent in core ciliome genes (present in 62% of this gene category) than in subtype specific ciliome (present in 6% of these genes) (**Figure 2F** and **Figure 2 Source data 2**). In addition, most X-boxes are located less than 600 bp from the center of its closest FKH-8 ChIP-seq peak (**Figure 2G**), this close proximity is a feature of core-ciliome components (found in 66% of core ciliome genes) but not for sub-type ciliome genes (only 9% of X-boxes found in this category contain a nearby FKH-8 peak) (**Figure 2 Figure Supplement 2** and **Figure 2 Source data 2**).

Finally, we assessed FKH-8 and DAF-19 physical interaction by co-expression in HEK293 cells and co-immunoprecipitation. Our results show that FKH-8 and DAF-19 interact bound to DNA (**Figure 2 Figure Supplement 2** and **Figure 2 Figure Supplement 2 Source data 1**) and more importantly also in the nuclear soluble fraction free of chromatin (**Figure 2H** and **Figure 2 source data 3**). In

summary, our data is consistent with FKH-8 and DAF-19 acting together to regulate ciliome gene expression, particularly in core cilia components.

fkh-8 mutants show defects in ciliome reporter gene expression

The only available *fkh-8* mutation, *tm292*, is a deletion downstream the FKH DNA binding domain, suggesting it might not be a null allele (**Figure 2A, Figure 2 Source data 1**). Thus, we built *fkh-8(vlc43),* a deletion allele that removes the whole *fkh-8* coding region (**Figure 2A**). We selected eight reporters for six genes that code for core cilia components and that overlap with FKH-8 ChIP-seq peaks (**Figure 1 Figure Supplement 1**) and analysed their expression in *fkh-8(tm292)* and *fkh-8(vlc43)* mutants.

Both *fkh-8* mutant alleles show significant expression defects in all reporters except for *tub-1*/Tub and the long *peli-1/*Peli1,2,3 and *xbx-1* reporters (**Figure 3A, B Figure 3 Figure Supplement 1 and Figure 3 Source data 1**). Lack of fluorescence reporter expression in *fkh-8* mutants reflects enhancer activity defects and not the absence of the ciliated neurons *per se*, as *tub-1*/Tub and the long *peli-1/*Peli1,2,3 reporters are expressed in 53 and 46 ciliated neurons respectively in *fkh-8* mutants, similar to *wild type* expression levels (**Figure 3 Figure Supplement 1 and Figure 3 Source data 1**). Phenotypes are often more penetrant in *fkh-8(vlc43)* null allele than in the *fkh-8(tm292)* and both *fkh-8(vlc43)* and *fkh-8(tm292)* heterozygote animals show similar reporter expression levels as *wild type* indicating both alleles are recessive and *tm292* is a hypomorph (**Figure 3 Figure Supplement 1**).

Endogenously tagged *osm-5* core cilome gene *[osm-5(syb6528), osm-5::SL2::GFP::H2B]* shows panciliary expression in *wild type* animals (**Figure 3C**),

fluorescence intensity is greatly reduced in *fkh-8(vlc43)* animals (**Figure 3C, D**) further supporting the role of FKH-8 in direct control of ciliome gene expression. *fkh-8(vlc43)* animals show missing *ift-20* expression in ten neurons including the four pairs of dopaminergic ciliated mechanosensory neurons (CEPV, CEPD, ADE and PDE). Expression in *fkh-8(vlc43)* animals of *fkh-8* cDNA under the control of a *dat-1* dopaminergic specific promoter, which is unaffected in *fkh-8* mutants (see below **Figure 7 and Figure 7 Source Data 1**), is able to rescue *ift-20* reporter expression, consistent with a cell autonomously role of *fkh-8* in the regulation of ciliome gene expression.

Next, we complemented the TF mutant analysis with *cis*-regulatory mutant analysis. We focused on *ift-20* and short *xbx-1* reporters which both overlap with FKH-8 ChIP-seq peaks (**Figure 3 Figure Supplement 2**). Three independent transgenic lines with point mutations for FKH binding sites show broad expression defects both for *ift-20* and *xbx-1* reporters (**Figure 3 Figure Supplement 2 and Figure 3 Source data 1, 2**). *cis*-mutation expression defects are stronger than the ones observed for *fkh-8* mutant alleles suggesting either other FKH factors can compensate the lack of *fkh-8* or that *cis*-mutations could affect the binding of other TFs in addition to FKH-8. Future work will be required to assess if other FKH TFs are expressed in specific subpopulations of ciliated neurons and if they can compensate for the lack of FKH-8.

In summary, our *cis* regulatory and *fkh-8* mutant analyses together with ChIP-seq data unravel a cell autonomous role for FKH-8 in the direct regulation of ciliome gene expression.

FKH-8 and DAF-19/RFX act synergistically

346 FKH-8 binds five different locations in the *daf-19* locus (**Figure 4A**) while *fkh-8*
347 locus contains 3 putative X-box sites (**Figure 2A**), suggesting they could cross-
348 regulate each other's expression. Transcription of *daf-19* generates different
349 isoforms that share the carboxyl terminal (Ct) domain and the RFX DNA binding
350 domain but differ in the amino-terminal region (**Figure 4A**). Some of these
351 isoforms are expressed in a mutually exclusive manner: *daf-19d* is specifically
352 expressed in ciliated neurons while *daf-19a/b* isoforms are expressed in the rest
353 of the nervous system but not in ciliated neurons (Senti and Swoboda, 2008).
354 Accordingly, a fosmid based Ct-tagged DAF-19 reporter that labels all isoforms
355 is broadly expressed in neurons (**Figure 4 Figure Supplement 1**). We did not
356 find any obvious DAF-19::GFP expression defects in *fkh-8(vlc43)* mutants
357 (**Figure 4 Figure Supplement 1**). Co-localization of DAF-19::GFP with *dat-*
358 *1::mcherry* dopaminergic reporter expression or DiD lipophilic staining also
359 reveals similar expression in *wild type* and *fkh-8(vlc43)* mutants in the
360 dopaminergic or amphid ciliated neurons (**Figure 4 Figure Supplement 1**). Thus,
361 our data suggest that, despite its extensive binding to *daf-19* locus, FKH-8 does
362 not seem to be required for *daf-19* expression in ciliated neurons.
363 Next, we assessed FKH-8::GFP fosmid expression in *daf-19(m86); daf-*
364 *12(sa204)* double mutant. In contrast to ciliated-neuron specific expression
365 pattern seen in *wild type,* FKH-8::GFP is expressed seemingly pan-neuronally in
366 *daf-19(m86); daf-12(sa204*) double mutants (**Figure 4B**), suggesting a
367 repressive role for DAF-19. FKH-8::GFP expression in the PDE dopaminergic
368 ciliated sensory neuron is unaffected in *daf-19(m86); daf-12(sa204*) double
369 mutants as assessed by co-localization with *dat-1:cherry* [91% PDE neurons are
370 FKH-8::GFP positive in *wild type* animals and 92% in *daf-19(m86)* mutants],

371 suggesting that FKH-8::GFP expression is unaffected by the lack of *daf-19* in
372 ciliated neurons (**Figure 4 Source data 1**). *daf-19(m86)* allele affects all isoforms;
373 as DAF-19 isoform D is expressed in ciliated neurons, our results suggest DAF-
374 19D is not necessary for FKH-8 expression in ciliated neurons. In contrast, DAF-
375 19 isoforms A and B seem to repress FKH-8 expression in non-ciliated neurons,
376 in agreement with this hypothesis, *daf-19(of5),* a mutant allele that specifically
377 affects *daf-19 a/b* isoform expression (**Figure 4A**) shows similar pan-neuronal
378 de-repression of FKH-8::GFP (**Figure 4B**). Consistent with the role for DAF-
379 19A/B as repressor (De Stasio et al., 2018; Senti and Swoboda, 2008), we found
380 also found *kap-1* ciliome gene reporter is de-repressed in *daf-19(m86); daf-*
381 *12(sa204)* mutants (**Figure 4 Figure Supplement 2**). Interestingly, both *kap-1*
382 expression in ciliated neurons as well as *kap-1* de-repression in *daf-19* mutants
383 is independent of FKH-8 (**Figure 4 Figure Supplement 2**). In AWB neurons, *kap-*
384 *1* reporter expression is regulated by FKH-2 transcription factor which is a
385 downstream target of DAF-19 (Mukhopadhyay et al., 2007). We find *fkh-2* fosmid
386 reporter (*wgIs185*) expression in AWB is also independent of FKH-8 (**Figure 4**
387 **Source data 1**) further supporting that *kap-1* expression in *wild type* and de-
388 repression in *daf-19* mutants is independent of FKH-8.
389 Altogether our data supports previous results on the repressive role of DAF-19
390 A/B long isoforms (De Stasio et al., 2018; Senti and Swoboda, 2008), which at
391 least in part seems to be independent of FKH-8 activity.
392 Next, we assessed sufficiency of DAF-19 and FKH-8 for the expression of *ift-20*
393 core ciliome gene reporter. Using a heat shock inducible promoter we ectopically
394 expressed DAF-19D during embryonic development. DAF-19D induction was
395 sufficient to produce ectopic *ift-20* expression while similar FKH-8 expression did

not produce any significant effect (**Figure 4 Figure Supplement 2**). These results, together with the stronger phenotypes of *daf-19(m86)* mutants, indicate that although both TFs are required for ciliome gene expression, RFX TFs display more instructive functions in ciliome gene expression.

Considering that DAF-19 and FKH-8 share several ciliome gene targets, bind the regulatory regions of ciliome genes in close proximity and both can physically interact, we aimed to assess if DAF-19 and FKH-8 cooperate in the regulation of ciliome gene expression. *daf-19* and *fkh-8* genes are both located in chromosome II, despite several attempts, we failed to generate *daf-19(m86), fkh-8(vlc43) II; daf-12(sa204*) triple mutants but we were able to obtain *daf-19(m86), fkh-8(tm292); daf-12(sa204*) recombinant animals. To study synergistic effects, we counted the number of reporter expressing cells in the different genetic backgrounds. Loss of *ift-20, peli-1* and *xbx-1* reporter expression in triple mutants is greater than what will be expected for the multiplicative effect of the individual *daf-19(m86); daf-12(sa204*) and *fkh-8(tm292); daf-12(sa204*) phenotypes (**Figure 4C, D and Figure 4 Figure Supplement 2**) suggesting FKH-8 and DAF-19 could act synergistically. Of note, these reporters still show some vestigial expression in the triple mutant (**Figure 4C and Figure 4 Figure Supplement 2**). We CRISPR-engineered a full deletion of the *fkh-8* locus in the *daf-19(m86); daf-12(sa204); ift-20::rfp* strain which generated a viable triple null mutant [*fkh-8 (vlc39)* allele]. These animals show similar residual *ift-20* expression in a couple neurons (**Figure 4C**), based on location we tentatively identified them as CEPV and BAG neurons, however lack of co-expression using specific reporters (*otIs259 [dat-1::gfp] and rpIs3 [gcy-33p::GFP])* discard these neuron identities. Taking into consideration that *daf-19(m86)* mutants show ectopic expression for

some ciliome genes and the absence of other ciliated neuron candidates in the region, we conclude *ift-20* positive cells in triple mutants are not likely to be ciliated neurons.

Importantly, although our analysis of total number of cells suggest synergy between FKH-8 and DAF-19, ectopic reporter expression in *daf-19(m86)* mutants could be masking additional synergistic effects. Thus, to unequivocally determine synergy, we labelled specific ciliated neuron populations in the different genetic backgrounds. We find *osm-1* reporter expression in CEPV and CEPD neurons is unaffected or only slightly affected in *fkh-8(tm292); daf-12(sa204) and daf-19(m86); daf-12(sa204)* double mutants compared to *daf-12(sa204)* controls but is completely abolished in triple mutants (**Figure 4E and Figure 4 Source data 1**). Similarly, *xbx-1* expression in PHA neurons is only significantly affected in triple mutants (**Figure 4E and Figure 4 Source data 1).** These data strongly suggest FKH-8 and DAF-19 act synergistically in the expression of specific ciliome genes and cellular contexts.

FKH-8 is required for correct cilia morphology

Mutations in several ciliome core components, including *osm-5* and *xbx-1*, whose reporters are affected in *fkh-8* mutants, show cilium morphology defects (Blacque et al., 2004; Mukhopadhyay et al., 2007; Perkins et al., 1986; Starich et al., 1995). In addition, cilia length is controlled by a balance between cilia assembly and disassembly regulated by IFT and mutants for ciliome components can produce both shortened or elongated cilia (Blacque et al., 2004; Burghoorn et al., 2007; Fujiwara et al., 1999). One of the most commonly used methods to assess gross cilium integrity is lipophilic dye staining (like DiD), which in *wild type* animals

446 labels a subpopulation of amphid and phasmid neurons (Starich et al., 1995).
447 Despite ciliome gene expression defects, *fkh-8(vlc43)* animals show similar DiD
448 staining compared to *wild type* (**Figure 4 Source data 1**).
449 Next, we directly analysed cilium morphology labelling specific subpopulations of
450 ciliated neurons (**Figure 5, Figure 5 Source data 1**). Cilium length in CEP and
451 AWB neurons is significantly reduced in *fkh-8(vlc43)* mutants compared to
452 controls, while ADF neuron cilium length is significantly increased (**Figure 5,**
453 **Figure 5 Source data 1**). In addition, *fkh-8* mutants display arborization defects
454 in AWA cilia (**Figure 5, Figure 5 Source data 1**). Of note, *osm-5* kinesin, which
455 our data show is a direct target for FKH-8, is required for correct AWB cilia
456 morphology (Mukhopadhyay et al., 2007), which is also affected in *fkh-8* mutants,
457 suggesting both phenotypes could be correlated.
458 To discard cilia morphology defects were caused by selective effects of sodium
459 azide in the mutants we treated worms with levamisole. Unexpectedly, we found
460 *fkh-8(vlc43)* mutants are resistant to this drug. Of note, *daf-19* mutants are also
461 levamisole resistant (Senti and Swoboda, 2008), which suggest this phenotype
462 for both mutants could be related to their ciliogenic functions. Alternatively, we
463 used polystyrene beads to physically immobilized worms. We found that cilia
464 morphology defects were also present in *fkh-8(vlc43)* mutants using this method
465 (**Figure 5 Figure Supplement 1**). Thus, FKH-8 is necessary to regulate correct
466 cilium length and morphology in diverse types of ciliated neurons.
467
468 ***fkh-8* mutants display defects in a wide range of cilia mediated behaviours**
469 In *C. elegans* cilia are necessary to mediate sensory functions (Bargmann, 1993);
470 thus, we interrogated *fkh-8* mutants with a battery of sensory paradigms.

fkh-8 mutants respond similarly to *wild type* animals to body gentle touch stimuli, which are mediated by not ciliated neurons (Chalfie and Sulston, 1981) (**Figure 6 Figure Supplement 1**), discarding general response or locomotory defects in *fkh-8* mutants. Response to posterior harsh touch, which is redundantly mediated by ciliated PDE and non-ciliated PVD neurons (Li et al., 2011a) is also unaffected in *fkh-8(tm292)* and *fkh-8(vlc43)* animals, suggesting FKH-8 is not required to mediate this mechanosensory behaviour (**Figure 6 Figure Supplement 1**).

We tested two additional mechanosensory behaviours mediated only by ciliated sensory neurons: basal slowing response, mediated by dopaminergic ciliated neurons (Sawin et al., 2000a) and nose touch, mediated by ASH, FLP and OLQ ciliated neurons (Kaplan and Horvitz, 1993a). No defects on basal slowing response are found in *fkh-8(vlc43)* null mutants, while both *fkh-8* alleles are defective for nose touch responses (**Figure 6A, B, Figure 6 Data source 1)**. *vlc43* null allele shows stronger defects than *tm292* allele, supporting the hypomorphic nature of *tm292* allele (**Figure 6A, Figure 6 Data source 1)**.

fkh-8(vlc43) animals are slightly but significantly dauer constitutive at 27ºC compared to N2 controls (**Figure 6C, Figure 6 Data source 1**), which might indicate *fkh-8* mutants show defects in preventing dauer entry, a process mediated by ADF, ASI and ASG ciliated neurons (Bargmann and Horvitz, 1991). Moreover, exposure to pheromones induces dauer entry in *fkh-8(vlc43)* animals less efficiently than in *wild type* animals [6 fold induction in *wild type* versus 3 fold induction in *fkh-8(vlc43)* animals] (**Figure 6C, Figure 6 Data source 1**), suggesting FKH-8 could also be required for correct dauer entry, which is mediated by ASJ ciliated neuron (Bargmann and Horvitz, 1991).

fkh-8(vlc43) null mutants, but not *tm292* allele, show significant odor avoidance defects to 2-nonanone (AWB mediated) and defective odor attraction to 2-heptanone (mediated by AWC) (**Figure 6D, E, Figure 6 Data source 1**) (Bargmann et al., 1993; Sengupta et al., 1996; Troemel et al., 1997a). Diacetyl attraction, which is mediated by AWA, is also decreased in *fkh-8(vlc43)* animals, although not significantly, due to high standard deviation values (**Figure 6F, Figure 6 Data source 1**).

Finally, we tested gustatory responses to NaCl, Sodium Dodecyl Sulfate (SDS) and copper. *fkh-8* mutants are attracted to NaCl similar to N2 controls, a response that is mediated mainly by ASE ciliated neurons (Bargmann and Horvitz, 1991) (**Figure 6G, Figure 6 Data source 1**). In contrast, avoidance response to SDS, mediated by ASH and ASK ciliated neurons (Hilliard et al., 2002) and avoidance to copper, mediated by ASH, ASE, ADF and ADL ciliated neurons (Guo et al., 2015; Sambongi et al., 1999), were significantly reduced both in *fkh-8(vlc43)* and *fkh-8(tm292)* animals (**Figure 6H, I, Figure 6 Data source 1**).

In summary, our battery of behavioural assays reveals FKH-8 is necessary for the correct response to a wide range of sensory stimuli (mechanical, gustatory or olfactory) that are mediated by different types of ciliated neurons (ADF, ADL, ASE, ASG, ASH, ASI, ASJ, ASK, AWB, AWC, FLP and OLQ). Some neurons controlling affected behaviours show corresponding morphological cilia defects in *fkh-8* mutants, such as ADF, AWA and AWB neurons. Nevertheless, we found that specific behaviours, such as attraction to NaCl or basal slowing response are sustained in *fkh-8* mutants, suggesting retained sensory functions for particular neuron types, even with gene expression or morphological cilia defects (such as CEPs).

Mouse FOXJ1 and FOXN4, master regulators of motile ciliome, can functionally replace FKH-8

Vertebrate FKH family is composed of 49 different members divided into 16 subfamilies (Shimeld et al., 2010). The establishment of specific orthology relationships between FKH members is challenging among distant species (Shimeld et al., 2010), precluding the direct assignment of the closest vertebrate ortholog for *C. elegans* FKH-8.

To date, no vertebrate FKH TF has been shown to be involved in ciliogenesis in primary cilia cell types. Nevertheless, in several vertebrate cell types that contain motile cilia, FoxJ1 FKH TF directly activates ciliome gene expression (Brody et al., 2000; Chen et al., 1998; Stubbs et al., 2008; Vij et al., 2012; Yu et al., 2008). Thus, considering its role in ciliogenesis, we next wondered if mouse FOXJ1 could functionally substitute FKH-8 in *C. elegans*. We find this to be the case as FoxJ1 cDNA expression under the dopaminergic promoter *dat-1* rescues *ift-20* expression similarly to *fkh-8* cDNA (**Figure 7A-C** and **Figure 7 Data source 1**). In *Xenopus*, another FKH TF, FoxN4, binds similar genomic regions to FoxJ1 and it is also required for direct ciliome gene expression in motile multiciliated cells (Campbell et al., 2016). We find FoxN4 expression also rescues *ift-20* expression defects in *fkh-8(vlc43)* animals. Importantly, we find that conserved functionality is not observed for any vertebrate FKH TFs as expression of mouse FoxI1, a FKH TF involved in the development of several tissues but not reported to control cilia gene expression (Edlund et al., 2015), does not rescue *fkh-8* mutant phenotype. In summary, our results unravel the functional conservation between FKH-8 and specific mouse members of the FKH family, which have already been described

to act together with RFX TFs in the regulation of ciliome gene expression in motile cilia cell types.

DISCUSSION

FKH-8 acts together with DAF-19 in the direct regulation of ciliome gene expression in sensory neurons

RFX are the only TFs known to be involved in the direct coregulation of ciliome gene expression both in cell types with motile and sensory cilia. This role is conserved in nematodes, flies and vertebrates (Choksi et al., 2014). In this work we characterized the persistent activity of ciliome gene reporters in *daf-19/RFX* null mutants, demonstrating that, in some specific cellular contexts, DAF-19/RFX is not necessary to drive ciliome gene expression. DAF-19 is the only RFX TF in *C. elegans*; thus, persistent enhancer activity must be attributed to other TF families.

A multi-angled approach allowed us to identify FKH-8 as a key regulator of ciliome gene expression in most, if not all, sensory neurons in *C. elegans*. FHK-8 is expressed almost exclusively in all ciliated neurons and binds to upstream regions of many ciliome genes. *fkh-8* mutants show decreased levels of ciliome reporter gene expression, abnormal cilia morphology and defects in a plethora of behaviours mediated by sensory ciliated neurons. Finally, mutations in putative FKH binding sites for two ciliome reporters lead to expression defects, further supporting the direct action of FKH-8 in ciliome gene expression. Altogether, our results show that FKH-8 has a key role in regulating ciliogenesis in sensory neurons and thus represents the first identified TF in any organism that acts together with RFX in cell types with non-motile primary cilia.

In the past, the identification of direct targets of RFX TFs has been instrumental in the identification of new ciliome components, which lead to a better understanding of cilia function and the etiology of ciliopathies (Blacque et al., 2005; Chen et al., 2006; Efimenko et al., 2005; Li et al., 2004; Schiebinger et al., 2019). FKH-8 binds to many genes in the *C. elegans* genome, some of them with uncharacterized functions; thus, similar to RFX, a more exhaustive characterization of FKH-8 targets could be used to uncover novel components of the sensory ciliome.

Specific DAF-19 isoforms repress *fkh-8* expression in non-ciliated neurons

Interestingly, our results show that DAF-19A and B isoforms repress (directly or indirectly) *fkh-8* expression in non-sensory neurons. Repression of alternative fates is a prevalent feature in neuronal development (Sousa & Flames, 2022). Repressive actions for DAF-19A/B have also been recently reported (De Stasio et al., 2018). Our results show that some ciliome components, such as *kap-1* can be upregulated in *daf-19* mutants even in the absence of FKH-8, thus DAF-19A/B repression of *fkh-8* might be necessary to avoid ectopic expression of some ciliome components but not others.

Role of FKH TFs in the transcriptional regulation of ciliome genes both in motile and sensory cilia cell types

Although TFs acting together with RFX in the regulation of ciliogenesis in sensory cell types were previously unknown, RFX TFs act in concert with the FKH TF FOXJ1 in the direct regulation of ciliome genes in different vertebrate cell types with motile cilia (Choksi et al., 2014).

Importantly, vertebrate sensory ciliogenesis is unaffected in FoxJ1 loss of function mutants (Choksi et al., 2014); thus, FoxJ1 role as a master regulator of ciliogenesis is restricted to motile ciliary cell types. In *Xenopus*, FoxN4 binds similar genomic regions to FoxJ1 and it is also required for motile ciliome gene expression (Campbell et al., 2016). We find both FOXJ1 and FOXN4, but not FOXI1, which has not been described to be involved in ciliogenesis, are able to functionally substitute FKH-8. In *C. elegans*, another FKH TF, FKH-2, is a downstream target of DAF-19 that controls expression of some ciliome components and cilium morphogenesis specifically in AWB neurons (Mukhopadhyay et al., 2007). Altogether, this data suggests that specific FKH TFs might have the capacity to act as direct ciliome regulators, independently of being expressed in motile or sensory cilia cell types.

FKH-8 and DAF-19 show synergistic actions

FKH-8 bound regions are enriched for X-box/RFX sites, FKH-8 and DAF-19 physically interact and double mutant analysis shows synergistic effects between *daf-19/RFX* and *fkh-8*, altogether these data suggest DAF-19 and FKH-8 cooperate in the regulation of a common set of regulatory regions. Cooperativity between FKH and RFX has been reported in vertebrates, in motile multiciliated cells of *Xenopus* larval skin FOXJ1 binding to ciliome gene promoters depends on the presence of RFX2 (Quigley and Kintner, 2017). In addition, in human airway multiciliated epithelial cell, RFX3 and FOXJ1 act synergistically in the activation of ciliome genes (Didon et al., 2013).

Importantly, cooperative actions of DAF-19 and FKH-8 seem prevalent for core ciliome components while several sub-type specific ciliome genes are known

621 targets for neuron type specific terminal selectors (Etchberger et al., 2007;
622 Flames and Hobert, 2009; Zhang et al., 2014).
623
624 **Evolution of cilia subtype specialization and ciliome regulatory logic**
625 Ancestral cilium present in the last common eukaryotic ancestor has been
626 proposed to combine motile and sensory functions (Mitchell, 2017). RFX role
627 regulating ciliome expression predates the emergence of metazoans, where
628 major cell type diversification has occurred (Chu et al., 2010; Piasecki et al.,
629 2010). FoxJ and FoxN constitute the most ancient FKH sub-families, present in
630 choanoflagelate *Monosiga brevicolis*, while FoxI subfamily is only present in
631 bilaterians (Shimeld et al., 2010). Moreover, the ability of RFX and FKH TFs to
632 bind similar genomic regions is not limited to metazoans and it is also present in
633 fungi. For example, in *Schizosaccharomyces pombe,* which lacks cilia and
634 ciliome genes, Fkh2 FKH TF and Sak1 RFX TF bind the same regulatory regions
635 to control cell cycle gene expression (Garg et al., 2015), suggesting that the joint
636 actions for these TFs could be present before the split of fungi and metazoans.
637 Alternatively, RFX and FKH TFs might have an inherent ability to cooperate that
638 could explain convergent evolution of these TFs in ciliome regulation both in
639 sensory and motile cilia cell types (Sorrells et al., 2018).
640 In light of these data, we hypothesize that RFX and FKH role as co-regulators of
641 ciliome gene expression could precede the emergence of cilia division of labor
642 and the specialization of motile and sensory cilium in different cell types (**Figure**
643 **7D**).
644
645 **Role of FKH TFs in ciliome regulation of primary cilia cell types**

Regardless of the evolutionary history of events underlying RFX and FKH functions as direct regulators of ciliome gene expression, our results raise the possibility that, in vertebrates, yet unidentified FKH TFs could act together with RFX in the regulation of ciliome gene expression in sensory ciliated cell types (**Figure 7D**). The establishment of specific orthology relationships between FKH members among distant species is challenging (Larroux et al., 2008; Shimeld et al., 2010) precluding the direct assignment of the closest vertebrate ortholog for *C. elegans* FKH-8. In addition, functional paralog substitutions among TFs of the same family have been described to occur in evolution (Tarashansky et al., 2021). Importantly, FoxJ1 and FoxN4 mutants do not show ciliome gene expression defects in non-motile ciliated cell types (Brody et al., 2000; Campbell et al., 2016; Chen et al., 1998; Stubbs et al., 2008; Yu et al., 2008). Other members of FoxJ and FoxN subfamilies are broadly expressed in mouse neurons, which all display primary cilia (Zeisel et al., 2018). It will be important, in future studies, to determine if additional FoxJ and FoxN TFs can rescue *fkh-8* expression defects in *C. elegans* and if they display similar roles in mammals. These studies could also help better characterize the functional meaning of non-coding mutations associated to orphan ciliopathies.

METHODS

Materials availability

Newly generated strains are listed in **Supplementary File 1** and accessible through Caenorhabditis Genetics Center (CGC). Plasmids are available upon request.

C. elegans strains and genetics

C. elegans culture and genetics were performed as previously described (Brenner, 1974). Strains used in this study are listed in **Supplementary File 1**.

Mutant strain genotyping

Mutant strains used in this study are listed in **Supplementary File 1**. Deletion alleles were genotyped by PCR. Presence of *daf-19(m86)* allele was determined by visual inspection of the dye-filling defective phenotype of homozygous mutants. Presence of *daf-12(sa204)* allele was ensured through a double cross strategy, crossing of F1 males with original *daf-12(sa204)X* mutants. Strains carrying point mutations were genotyped by sequencing. Genotyping primers are included in **Supplementary File 1**.

DiD staining.

Lipophilic dye filling assays were performed with the 1,1′-dioctadecyl-3,3,3′,3′-tetramethylindodicarbocyanine, 4-chlorobenzenesulfonate salt (DiD) (Thermofisher, #D7757). DiD staining solution was freshly prepared prior to every assay as a 1:200 dilution of the DiD stock solution [2 mg/mL dilution in N,N-dimethyl formamide (Sigma, #D4551)] in M9 1X buffer. Animals were transferred into 1.5 mL tubes containing 200 μL of the DiD staining solution and incubated (wrapped in aluminium foil) for 2 hours at room temperature in an orbital shaker in a horizontal position. Animals were collected with a glass Pasteur and transferred to fresh NGM plates. Robust identification of the ASK, ADL, ASI, AWB, ASH, ASJ, PHA and PHB ciliated neurons was achieved through this method.

Generation of *C. elegans* transgenic lines

Fluorescent reporters for ciliome genes were generated through fusion PCR (Hobert, 2002). To facilitate identification and scoring of reporter-expressing cells,

GFP was tagged to the cell's nucleus employing a modified sequence of the classical SV40 large T antigen nuclear localizing signal (NLS) (Kalderon et al., 1984). Regulatory sequences were amplified with custom oligonucleotides from N2 genomic DNA preparations. An independent PCR was used to amplify the 2xNLS::GFP::*unc-54* 3'UTR fragment from an NLS version of the pPD95·75 plasmid (pNF400). Successfully fused PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN, #28106) and resuspended in nuclease-free water (Sigma, #W4502).

Mutated versions for the *xbx-1* and *ift-20* promoters were generated as PCR products by introducing the desired mutation of putative FKH sites within the corresponding custom primers. Putative FKH sites were identified through the single sequence scan tool from the CIS-BP website (Weirauch et al., 2014). Mutation criteria accounted for the nature of the nitrogenous bases and the number of hydrogen bonds they could form; thus, A was mutated to C and G was mutated to T (and vice versa). Mutated sequences were checked to discard the generation of new TF binding site motifs using both the motif scan tool of the CIS-BP database and the Tomtom tool (Gupta et al., 2007) from the MEME Suite website. When designed mutations created potential new TF binding sites manual punctual mutations were applied to disrupt those potential sites.

To generate FKH-8 rescuing plasmids, constructs containing the cDNA of the corresponding FKH TF fused to the self-cleaving peptide T2A (Ahier and Jarriault, 2014) and the eGFP cDNA from the pPD95.75 plasmid were created. Such constructs were then cloned under the control of the dopaminergic *dat-1* promoter between the KpnI/XhoI sites of the pPD95.75 backbone vector. *fkh-8* cDNA sequence was synthetically generated (Biomatik). Murine FKH members were

obtained as Dharmacon clones (FoxJ1: MMM1013-202732974, FoxN4: MMM1013-211694291, FoxI1: MMM1013-202763055).

Simple-array transgenic lines were generated by intragonadal microinjection into strains of the appropriated genotype. The injection mix was composed by 50 ng/µL of a given purified fusion PCR or a rescuing plasmid plus 100 ng/µL of the pFR4 plasmid, *rol-6(su1006),* as a co-marker (Mello et al., 1991).

Generation of *C. elegans* mutations

Whole deletion of the *fkh-8* locus was performed through a co-CRISPR strategy (Kim et al., 2014) using *dpy-10(cn64)* as conversion marker (Arribere et al., 2014). Custom CRISPR RNAs (crRNAs) were ordered (IDT, Alt-R® CRISPR-Cas9 crRNA XT) targeting both sides of the desired deletion of *fkh-8* and at the 5' of the *dpy-10* site of mutation. Single stranded oligodeoxynucleotide (ssODNs) of approximately 100 base pairs overlapping each side of the genetic modifications were also ordered (IDT) and used as donor templates to achieve homology-directed repair. Cas9 nuclease (IDT, #1081058) and the universal trans-activating crRNA (tracrRNA) needed to initiate enzymatic activity (IDT, #1072532) were used. co-CRISPR injections were performed on young adult hermaphrodites expressing the reporter *otIs395(ift-20::NLS::tagRFP)III*. Microinjection mix was freshly prepared with all 3 crRNAs plus the tracrRNA, ssODNs and Cas9 nuclease. Ribonucleoprotein complex formation was achieved by incubating this mix for 10 minutes at 37 Celsius degrees. Before use, the final mix was incubated on ice for 30 minutes. *daf-19(of5)* was generated by CRISPR deletion of 9 bp that include the ATG for the *daf-19 a* and *daf-19 b* isoforms and inserting a guanine nucleotide (N2 sequence: GGCAGAGAAGAAAGT**CATGACCA**ATGAGGAGCC; of5 sequence

GGCAGAGAAGAAAG**g**ATGAGGAGCC. Knock-in strain PHX6528 *[osm-5(syb6528), osm-5::SL2::GFP::H2B]* was generated by SunyBiotech's CRISPR services. All custom primer sequences and concentrations used for the generation of the aforementioned strains are included in the **Supplementary File 1**.

Behavioural assays

Unless otherwise stated, all mechano- and chemosensory assays were performed over small-scale synchronized populations of young adult hermaphrodites.

Nose touch tests were performed as previously described (Kaplan and Horvitz, 1993b). Ten minutes before the assay, young adult hermaphrodites were transferred to non-seeded NGM agar plates and nose touch responses were elicited by causing a nose-on collision placing an eyelash attached to a pipette tip in the path of an animal moving forward. With brief modifications from (Brockie et al., 2001), five consecutive nose touch trials were scored for each worm.

Both gentle and harsh touch mechanosensory tests were performed as previously described (Chalfie et al., 1985). Briefly, gentle touch assays were performed by alternatively stroking the animal just behind the pharynx and just before the anus with an eyebrow hair attached to a pipette tip for a total amount of 10 strokes (Hobert et al., 1999). Harsh touch assays were also performed by stroking the worms across the posterior half of their bodies in a top-down manner with a platinum wire. Each worm was tested five times with a 2 minutes interval between each trial (Li et al., 2011).

For all aforementioned mechanosensory assays, escape responses of trailed animals were recorded and a population response index (RI) was calculated for every replica as: RI = total number of escape responses / total amount of strokes Chemotaxis towards diacetyl, 2-heptanone, NaCl and 2-nonanone were performed over 3 times freshly washed worms with 1 mL of filtered, autoclaved CTX solution, aspirating the supernatant to a final volume of approximately 100 µL. 2 µL of this worm-containing solution were placed at the proper place of the assay plates. During the assays, worms were allowed to freely crawl across the plates for 60 minutes at room temperature and then stored at 4 °C until the next day when worms' positions were scored and behavioural indexes were calculated.

With few modifications, volatile diacetyl attraction assay was performed as described by (Margie et al., 2013a). A four-quadrant paradigm drawn at the base of non-seeded NGM agar plates was used, adding a 1 cm circular central area that worms had to trespass to be scored. Stock diacetyl (Sigma-Aldrich, #803528) test solution was prepared as a 0.5% V/V mix in absolute ethanol (Scharlau, #ET00101000). Absolute ethanol was used as control solution. Immediately after the worms were plated, 2 µL of a mix combining equal volumes of diacetyl stock solution and sodium azide 1M were pipetted onto the 2 test sites (T) of the agar plate. Same procedure was then performed for the 2 control sites (C). Chemotaxis index (CI) was then calculated as: CI = (worms in (T1 + T2) - worms in (C1 + C2)) / total scored worms

Chemotaxis assay towards 2-heptanone was performed as previously reported (Zhang et al., 2016). A two-halves paradigm was used, adding the threshold distance by (Margie et al., 2013b) to prevent immobile worms from skewing the

795 data. 2-heptanone (Sigma Aldrich, #W254401) test solution was prepared as a
796 1:10 V/V mix in ethanol absolute. Ethanol was used as control solution.
797 Immediately after the worms were plated, 3 µL of a mix combining equal volumes
798 of 2-heptanone stock solution and sodium azide 1M were pipetted onto the test
799 site (T) of the agar plate. Same procedure was follow to the control site (C). CI
800 was calculated: CI = (worms in T - worms in C) / total scored worms.
801 Chemotaxis toward NaCl was also performed of a two-halves paradigm. Radial
802 gradients of either test or control solutions were created prior to worm loading as
803 originally stated (Ward 1973). Following (Frøkjær-Jensen et al., 2008), 10 µL of
804 NaCl (Sigma, #S3014-1KG) 2.5 M (dissolved in double distilled water (ddH2O))
805 or ddH2O itself were respectively pipetted onto the agar surface at T and C spots
806 and allowed to diffuse for 12-14 hours at room temperature. To increase
807 steepness of the gradients, 4 µL of NaCl 2.5 M or ddH_2O solutions were
808 additionally added to the T and C spots respectively 4 hours prior to the
809 chemotaxis assay. Chemotaxis indexes for two-halves paradigm assays were
810 calculated as: CI = (worms in T - worms in C) / total scored worms.
811 Avoidance assay towards 2-nonanone was performed as previously reported
812 (Troemel et al., 1997). Briefly, six equal sectors labelled as A, B, C, D, E and F
813 were drawn on the base of squared plates (90 x 15 mm, Simport™, # 11690950)
814 containing 15 mL of standard NGM agar. Stock 2-nonanone (Sigma-Aldrich,
815 #W278550) test solution was prepared as a 1:10 V/V mix in absolute ethanol.
816 Ethanol was used as control solution. Immediately after the worms were plated
817 on the centre of the plate, 2 µL of a mix combining equal volumes of 2-nonanone
818 stock solution and sodium azide 1M were pipetted onto two spots within
819 peripheral test sector A. Same procedure was then performed for the ethanol

control sites within sector opposite peripheral control sector F. Population avoidance index (AI) was calculated as: AI = (worms in (A+B) - worms in (E + F)) / total amount of worms.

Avoidance responses to water-soluble compounds were evaluated using the drop test as previously described (Hilliard et al., 2004) following a few modifications. Well-fed synchronized young adult hermaphrodites were washed three times with M13 buffer, 5 animals were then placed on unseeded NGM agar plates and allowed to rest for 10 minutes. Two test solutions were assayed: 0.1% W/V sodium dodecyl sulfate (SDS) (Sigma, #L3771-100G) and 0.1 mM CuSO4 pentahydrate (Merck, #1027901000), both dissolved in the M13 buffer that acted as control solution. Each animal was tested first with 4 single drops of the control solution and then with 4 single drops of the testing solution, allowing for 2 minutes of recovery between each stimulus. Avoidance response was scored within 4 seconds after substance delivery. Population avoidance index (AI) per genotype and replica was calculated as: AI = number of responses / total amount of drops.

Dauer induction was performed using filtered liquid culture obtained from *wild type* worms grown at 7 worms/µl for 4 days. Briefly, 300µl of pheromone containing extracts or control extracts (culture media alone) were added to 60mm OP50-seeded NGM plates. After drying, 10 gravid worms were added and allowed to lay eggs for 18 hours and then removed from the plates. 72h later, resulting P0 worms were scored and percentage of dauer animals determined for each condition. Dauer induction was carried at 27ºC in four independent experiments performed in parallel with wild type and *fkh-8(vlc43)* mutant worms.

Basal slowing response was performed with few modifications as previously reported (Sawin et al., 2000). In this case, 60 mm NGM plates in which HB101

was seeded in only one half of the plate were used. Briefly, well-feed worms were 3 times washed with 1 mL of filtered, autoclaved CTX solution, supernatant aspirated to a final volume of approximately 200 µL and 2 µL of this worm-containing solution (with no less than 10 animals) was placed at the non-seed part of pre-warmed assay plates. Free movement of the worms across the plates was recorded capturing 30 frames per second. Body beds per 20 seconds intervals were counted from same worms moving on agar and crawling across the bacterial lawn.

Sample size, tested genotypes, number of animals and number of replicates performed per assay are detailed in the corresponding figure legends and in Source Data Files. All strains used for these behavioural studies are listed in **Supplementary File 1.**

Co-immunoprecipitation experiments

Human optimized sequences for FLAG:DAF-19 and HA:FKH-8 were cloned into PCDNA3.1 plasmid. A 10cm plate of HEK293 cells (Human 293T, ATCC CRL-3216, authenticated by microsatellite amplification, Secugen, tested micoplasm negative) was transfected with 5µg of each plasmid. At 48h post-transfection cells were washed three times with ice-cold PBS and lysed for 10 min on ice with 200 µl of cytoplasmic fractionation buffer containing 10mM Hepes pH 7.9; 10mM KCl; 1.5mM MgCl2; 0.34M sucrose, 10%glycerol; 1mM DTT; 5ug/ml protease inhibitor; 0.1mM PSMF and 0.1% Triton X-100 from a 10% stock. Cells were centrifuged at 3500 rpm for 4 min at 4°C. The supernatant (cytoplasmic fraction), was collected and clarified by centrifugation (15 min 14.000 rpm at 4°C).

The pellet (nuclei fraction) was lysed for 30 min on ice with 100 µl of nuclear fractionation buffer comprising 10mM Hepes pH 7.9, 3 mM EDTA, 0.5 mM NaF,

0.2 mM EGTA, 1mM DTT; 5ug/ml protease inhibitor and 0.1mM PSMF. The fraction was homogenized, incubated for 10 min on ice and centrifuged at 4000 rpm for 4 min at 4°C. The supernatant (soluble nuclear fraction), was collected. The pellet (chromatin fraction) was resuspended with 300 µl of nuclear fractionation buffer and was sonicated for 15-20 seconds (25% amplitude).

Total protein concentration of soluble cytoplasmic fraction, soluble nuclear fraction and chromatin fraction were used for coimmunoprecipitation assays. The association between DAF-19 and FKH-8 was analyzed by coimmunoprecipitation using anti-HA magnetic beads (bimake.com). Cytoplasmic and nuclear fractions were incubated with 10 µl of magnetic beads overnight at 4ºC with rotation. The beads were separated by a magnetic separator and washed 3 times in 500µl TNE buffer 0.1% Triton X-100. Finally, for reducing SDS-PAGE analysis, 30µl 2xLaemmli buffer (with 2% ß-MeOH) was added and the samples were boiled for 5 minutes at 96º C. Anti-mouse Light chain-specific secondary antibodies (Jackson Immunoresearch ref: 115-655-174) were used for detection since the heavy chain from the immunoprecipitation would mask the FKH-8 signal. anti-HA antibody (Biolegend 901501), anti-FLAG antibody (Sigma, F1804). Cytoplasmic and nuclear fractions were controlled using antibodies against MEK2 (Becton Dickinson, 610235), marker for cytoplasmic fraction and histone H3, marker for chromatin associated fraction (Abcam ab1791).

Microscopy

For scoring and image acquisition, worms were anesthetized with a drop of 0.5 M sodium azide (Sigma, #26628-22-8) on 4% agarose pads (diluted in distilled water) placed over a regular microscope glass slide (Rogo Sampaic,

#11854782). These preparations were sealed with a 24 x 60 mm coverslip (RS France, #BPD025) and animals were then conveniently examined.

Scoring of ciliome features was performed observing the animals on a Zeiss Axioplan 2 microscope using a 63X objective. Assessment of fluorescence signal on PDE and Phasmid regions was performed *de visu*. To appropriately assess number of cells in the head, optical sections containing the volume of reporter-positive neurons in the head of the animals were acquired at 1 µm intervals and images were manually scored using FIJI (Schindelin et al., 2012). Reporters used in the FKH cis-mutational analyses (both *wild type* and mutated forms) were scored *de visu* as the low intensity and fast bleaching in their signals precluded us from taking pictures.

Fluorescence intensity levels from the endogenous *osm-5::SL2::GFP::H2B* reporter strain were measured on young adult animals grown at 25 ºC. All images were acquired with a TCS-SP8 Leica Microsystems confocal microscope using a 63X objective on animals immobilised as previously described. Image acquisition was optimized considering the appropriate no saturating conditions for the *wild type* background. To avoid a possible bias induced by the volume of the worm's body, *dat-1::cherry* reporter was used to select a single section for the PDE, CEPV, CEPD and ADE nuclei and was used for *osm-5::GFP* fluorescence quantification. The contour of each nucleus was delineated and fluorescence intensity quantified using FIJI. For each cell, corrected total cell fluorescence was calculated as follows: Integrated Density – (Area of selected cell X Mean fluorescence of background reading).

For cilia morphology assessment, 0.5 M sodium azide was used as an immobilization agent image acquisition was performed with a TCS-SP8 Leica

Microsystems confocal microscope using a 63X objective. The following conditions of optical sections (µm) were used: CEP: 0.4 µm; ADF: 0.2; AWB: 0.24; AWC: 0.3. Retrieved images were z-projected at maximum intensity (Leica LAS X LS) and linear adjustment for brightness and contrast was performed prior to cilia length quantification (N ≥ 32 cilia per neuron type) (FIJI). AWA analysis was performed from images acquired from dorsoventrally positioned animals (N = 7) in which both cilia were levelled and depth of arborisation was estimated from the volume containing all the optical sections (0.3 µm) in which fluorescence signal was observed. Ciliary morphology was also measured on animals immobilised as previously reported (Niwa, 2017) using 10% (w/v) agarose pads and 2.5% solids (w/v) aqueous suspension of polystyrene microspheres with 100 nm of diameter (Polysciences, #00876-15). Optical sections containing the volume of reporter-positive cilia were acquired at 0.3 µm intervals. Images were z-projected at maximum intensity (Leica LAS X LS) and linear adjustment for brightness and contrast was performed prior to ciliary length quantification (N ≥ 30 cilia per neuron type) (FIJI).

All micrographs presented in this paper were acquired with a TCS-SP8 Leica Microsystems confocal microscope using a 63X objective and appropriate zooming conditions.

Statistical analyses.

Statistical significance for the mean number of reporter-positive neurons in whole animals among different genetic backgrounds was assessed by two-tailed t-test. Inbuilt Excel functions F.TEST and T.TEST were used and obtained p-values

were adjusted through Bonferroni correction accounting for all possible pairwise comparisons in each experiment.

To increase for statistical power, statistical significance for the mean number of reporter-positive neurons in the five distinct anatomical regions containing ciliated neurons among different genetic backgrounds was assessed by one-tailed t-test. Obtained p-values were then adjusted through the Benjamini-Hochberg procedure setting α level at 0.05. This same procedure was used to assess for statistical significance within the dauer induction experiments.

Unless otherwise stated, same two-tailed t-test procedure was followed in the assessment of statistical significance in behavioural experiment. Behavioural responses were ultimately analysed through the corresponding indexes ranging from 0 to 1 (or to -1 to 0 when avoidance responses were assayed). For each type of assay, a population-based mean index was calculated per replica and a final response index was then obtained as the mean of all replicas' means. Prior to hypothesis testing, the Shapiro-Wilk test (Shapiro and Martin, 1965) was used to address for the normality of these final indexes.

Assessment of synergistic effects between *fkh-8* and *daf-19* was performed under the multiplicative model (Wagner, 2015). Briefly, average number of reporter-expressing neurons found in the whole animals for each genetic background was transformed into the corresponding fold change related to the observed mean value in the *daf-12* single mutant. Next, expected values for the fold change corresponding to triple *daf-12*; *daf-19, fkh-8* mutants were calculated as the product of the mean observed values for the double *daf-12*; *daf-19* and daf-12; *fkh-8* mutant strains. To calculate the associated error for this indirect

968 measure, propagation error was used (sum of each standard deviation between
969 the corresponding mean from each observed value in each double mutant
970 multiplied by the value of the expected value). Statistical significance between
971 observed and expected values was then assessed through a one-sample t test.
972 For the assessment of statistical significance in rescue experiments, data was
973 categorically classified as ‘on’ or ‘off’ and the significance of the association was
974 examined using the two-tailed Fisher’s exact test. No further multiple testing
975 correction was performed, as *fkh-8* null mutants were exclusively compared to
976 wild type worms whereas each rescued line was exclusively compared against
977 the *fkh-8* null mutants.

978 **Bioinformatics Analysis**

979 Ciliome gene list was assembled including genes associated with cilium-related
980 terms from the Gene Ontology using AmiGO (Carbon et al., 2009), known ciliome
981 genes with functional X-boxes (Burghoorn et al. 2012) and genes whose
982 expression in ciliated neurons was reported in the WormBase. Transcription
983 factors were deliberately excluded from this list. A further curation process was
984 performed through a bibliographic research (see Figure 1 Source Data 1 for
985 complete ciliome gene list).
986 For each isoform of the final 163 genes composing the ciliome gene list, putative
987 regulatory sequences were retrieved from the Ensembl BioMart site (Kinsella et
988 al., 2011) spanning 700 base pairs in length upstream of their translational start
989 sites. These sequences were used to feed the RSAT oligo-analysis tool as
990 previously described (Defrance et al., 2008; Turatsinze et al., 2008), using as a
991 background model the in-tool genome of *C. elegans* and overall default options.
992 Retrieved matrices were then compared both against the CIS-BP 1.02 (Weirauch

et al., 2014) and the JASPAR core non-redundant 2018 (Khan et al., 2018) databases using the TomTom (Gupta et al., 2007) tool from the MEME suite (Bailey et al., 2009). Four different matrices of different lengths all matching RFX binding sites (X-boxes) were retrieved through this method. Overlapping of these matrices over the putative regulatory sequences from the ciliome genes was used to defined X-box regions whose coordinates were assessed for the ce10 version of the *C. elegans* genome.

Identification of candidate transcription factors with enriched expression in ciliated neurons was performed through the on-line tool GExplore$_{1.4}$ (Hutter and Suh, 2016), employing the sci-RNA-seq dataset by (Cao et al., 2017). A 5-fold enrichment ratio and a false detection rate of 0.001 were used.

Expression pattern data in each ciliated neuron type for candidate transcription factors at the fourth larval stage were retrieved from the *C. elegans* Neuronal Gene Expression Network (CeNGEN) (Taylor et al., 2021), whose results are freely accessible through the on-line tool SCeNGEA. Unfiltered data was used.

ChIP-seq data from *C. elegans* TFs were retrieved from the ENCODE portal website (Davis et al., 2018) (time of consulting: January the 10th, 2019). Peak annotation was carried out employing the ChIPseeker package (Yu et al., 2015), setting parameters as following: annotatePeak(gr1, tssRegion=c(-2000, 1000), level=lev, TxDb=annoData, overlap="TSS"). ENCODE accession numbers for all datasets used in this analysis are listed in Figure 1 Source Data 2.

fkh-8 ChIP-seq bed narrowPeak file (ENCODE accession: ENCFF653QKE) was used as input file for the web-based analysis tool ChIPseek (Chen et al., 2014). For de novo motif discovery, resulting fasta file with annotated peaks was then used to feed the RSAT peak-motifs tool as previously described (Thomas-

Chollier et al., 2012b, 2012a), setting the number of motifs per algorithm at 10 and using all 4 available discovery algorithms with overall default options.

For gene onthology, genes associated to FKH-8 ChIP-seq peaks where analysed through the on-line tool WormEnricher (Kuleshov et al., 2016).

Gene expression correlation between TFs and genes of interest were calculated using embryonic sc-RNA-seq data (Packer et al., 2019). Genes of interest were categorized into four categories: 1) core ciliome genes, 2) subtype-specific ciliome genes (both extracted from our ciliome list), 3) panneuronal genes (Stefanakis et al., 2015) and 4) ubiquitously expressed genes (Packer et al., 2019). In addition to fkh-8 and daf-19, the proneural TF factor hlh-14 was added as control TF not related to ciliogenesis. For all 10,775 ciliated cells present in the dataset, correlation index (R) between the expression levels for each gene and the TF was calculated. R data for each gene category are represented in the graph (See Figure 2 Source data 2 for R values).

Presence of RFX/*daf-19* binding motifs within the FKH TFs ChIP-seq peak sequences was performed with the on-line tool Centrimo (Bailey and MacHanick, 2012) from the MEME suite. Sequences 420 base pairs in length spanning 210 base pairs from the centre of each peak were extracted to prevent Centrimo from discarding sequences due to uneven sequence length within and among the different ChIP-seq datasets. This consensus length was used considering the average sequence length of FKH-8 ChIP-seq peaks. ENCODE accession numbers for all datasets used in this analysis are listed in Figure 2 Source data 2.

Visualization and analysis of ChIP-seq and RNA-seq files were performed with the Integrative Genomics Viewer (IGV) software (Robinson et al., 2011).

ACKNOWLEDGMENTS

We thank CGC (P40 OD010440) for providing strains. Dr Laura Chirivella, Noemi Daroqui, Anna Roig and Elia García for technical help. Erick Sousa for providing bioinformatics assistance. Ioannis Segos and Barbara Conrad for sharing the immobilization protocol, Ethel Queralt for advice on the Co-IP experiments and Ines Carrera and Elisa Martí for comments on the manuscript. Funding: This work was supported by European Research Council (StG2011-281920 and COG-101002203), Ministerio de Ciencia e Innovación (SAF2017-84790-R and PID2020-115635RB-I00) and Generalitat Valenciana (PROMETEO/2018/055).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Ahier A, Jarriault S. 2014. Simultaneous expression of multiple proteins under a single promoter in Caenorhabditis elegans via a versatile 2A-based toolkit. *Genetics* **196**:605–613. doi:10.1534/genetics.113.160846

Andreu-Cervera A, Catala M, Schneider-Maunoury S. 2021. Cilia, ciliopathies and hedgehog-related forebrain developmental disorders. *Neurobiol Dis* **150**:105236. doi:10.1016/j.nbd.2020.105236

Arribere JA, Bell RT, Fu BXH, Artiles KL, Hartman PS, Fire AZ. 2014. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in caenorhabditis elegans. *Genetics* **198**:837–846. doi:10.1534/genetics.114.169730

Ashique AM, Choe Y, Karlen M, May SR, Phamluong K, Solloway MJ, Ericson J, Peterson AS. 2009. The Rfx4 Transcription Factor Modulates Shh Signaling by Regional Control of Ciliogenesis. *Sci Signal* **2**(95). doi:10.1126/scisignal.2000602

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res*. Jul;37(Web Server issue):W202-8. doi:10.1093/nar/gkp335

Bailey TL, MacHanick P. 2012. Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res*. Sep 1;40(17):e128. doi:10.1093/nar/gks433

Bargmann CI. 1993. Genetic and cellular analysis of behavior in C. elegans. *Annu Rev Neurosci* **16**:47–71. doi:10.1146/annurev.ne.16.030193.000403

Bargmann CI, Hartwieg E, Horvitz HR. 1993. Odorant-selective genes and neurons mediate olfaction in C. elegans. *Cell* **74**:515–527. doi:10.1016/0092-8674(93)80053-H

Bargmann CI, Horvitz HR. 1991. Control of larval development by chemosensory neurons in Caenorhabditis elegans. *Science (1979)* **251**:1243–1246. doi:10.1126/science.2006412

Blacque OE, Perens EA, Boroevich KA, Inglis PN, Li C, Warner A, Khattra J, Holt RA, Ou G, Mah AK, McKay SJ, Huang P, Swoboda P, Jones SJM, Marra MA, Baillie DL, Moerman DG, Shaham S, Leroux MR. 2005. Functional genomics of the cilium, a sensory organelle. *Current Biology* **15**:935–941. doi:10.1016/j.cub.2005.04.059

Blacque OE, Reardon MJ, Li C, McCarthy J, Mahjoub MR, Ansley SJ, Badano JL, Mah AK, Beales PL, Davidson WS, Johnsen RC, Audeh M, Plasterk RHA, Baillie DL, Katsanis N, Quarmby LM, Wicks SR, Leroux MR. 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. doi:10.1101/gad.1194004

Bonnafe E, Touka M, AitLounis A, Baas D, Barras E, Ucla C, Moreau A, Flamant F, Dubruille R, Couble P, Collignon J, Durand B, Reith W. 2004. The Transcription Factor RFX3 Directs Nodal Cilium Development and Left-Right Asymmetry Specification. *Mol Cell Biol* **24**:4417–4427. doi:10.1128/mcb.24.10.4417-4427.2004

Brenner S. 1974. The genetics of *Caenorabditis elegans*. *Genetics* **77**:71–94. doi:10.1016/S0047-2484(78)80101-8

Brockie PJ, Mellem JE, Hills T, Madsen DM, Maricq A V. 2001. The C. elegans glutamate receptor subunit NMR-1 is required for slow NMDA-activated currents that regulate reversal frequency during locomotion. *Neuron* **31**:617–630. doi:10.1016/S0896-6273(01)00394-4

Brody SL, Yan XH, Wuerffel MK, Song SK, Shapiro SD. 2000. Ciliogenesis and left-right axis defects in forkhead factor HFH-4-null mice. *Am J Respir Cell Mol Biol* **23**:45–51. doi:10.1165/ajrcmb.23.1.4070

Burghoorn J, Dekkers MPJ, Rademakers S, de Jong T, Willemsen R, 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*. *Proceedings of the National Academy of Sciences* **104**:7157–7162. doi:10.1073/pnas.0606974104

Burghoorn J, Piasecki BP, Crona F, Phirke P, Jeppsson KE, Swoboda P. 2012. The in vivo dissection of direct RFX-target gene promoters in C. elegans reveals a novel cis-regulatory element, the C-box. *Dev Biol* **368**:415–426. doi:10.1016/j.ydbio.2012.05.033

1121 Campbell EP, Quigley IK, Kintner C. 2016. Foxn4 promotes gene expression
1122 required for the formation of multiple motile cilia. *Development* **143**:4654–
1123 4664. doi:10.1242/dev.143859
1124 Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C,
1125 Furlan SN, Steemers FJ, Adey A, Waterston RH, Trapnell C, Shendure J.
1126 2017. Comprehensive single-cell transcriptional profiling of a multicellular
1127 organism. *Science, Aug 18;357(6352):661-667*.
1128 doi:10.1126/science.aam8940
1129 Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, Lomax J, Mungall
1130 C, Hitz B, Balakrishnan R, Dolan M, Wood V, Hong E, Gaudet P. 2009.
1131 AmiGO: Online access to ontology and annotation data. *Bioinformatics*. Jan
1132 15;25(2):288-9. doi:10.1093/bioinformatics/btn615
1133 Chalfie M, Sulston J. 1981. Developmental genetics of the mechanosensory
1134 neurons of Caenorhabditis elegans. *Dev Biol* **82**:358–370.
1135 doi:10.1016/0012-1606(81)90459-0
1136 Chalfie M, Sulston JE, White JG, Southgate E, Thomson JN, Brenner S. 1985.
1137 The neural circuit for touch sensitivity in Caenorhabditis elegans. *Journal of*
1138 *Neuroscience* **5**:956–964. doi:10.1523/jneurosci.05-04-00956.1985
1139 Chatterjee S, Ahituv N. 2017. Gene Regulatory Elements, Major Drivers of
1140 Human Disease. *Annu Rev Genomics Hum Genet* **18**:45–63.
1141 doi:10.1146/annurev-genom-091416-035537
1142 Chen J, Knowles HJ, Hebert JL, Hackett BP. 1998. Mutation of the mouse
1143 hepatocyte nuclear factor/forkhead homologue 4 gene results in an
1144 absence of cilia and random left-right asymmetry. *Journal of Clinical*
1145 *Investigation* **102**:1077–1082. doi:10.1172/JCI4786
1146 Chen N, Mah A, Blacque OE, Chu J, Phgora K, Bakhoum MW, Newbury CRH,
1147 Khattra J, Chan S, Go A, Efimenko E, Johnsen R, Phirke P, Swoboda P,
1148 Marra M, Moerman DG, Leroux MR, Baillie DL, Stein LD. 2006.
1149 Identification of ciliary and ciliopathy genes in Caenorhabditis elegans
1150 through comparative genomics. *Genome Biol* **7**:R126. doi:10.1186/gb-
1151 2006-7-12-r126
1152 Chen TW, Li HP, Lee CC, Gan RC, Huang PJ, Wu TH, Lee CY, Chang YF,
1153 Tang P. 2014. ChIPseek, a web-based analysis tool for ChIP data. *BMC*
1154 *Genomics* **15**. doi:10.1186/1471-2164-15-539
1155 Choksi SP, Lauter G, Swoboda P, Roy S. 2014. Switching on cilia:
1156 transcriptional networks regulating ciliogenesis. *Development* **141**:1427–
1157 1441. doi:10.1242/dev.074666
1158 Chu JSC, Baillie DL, Chen N. 2010. Convergent evolution of RFX transcription
1159 factors and ciliary genes predated the origin of metazoans. *BMC Evol Biol*
1160 **10**:130. doi:10.1186/1471-2148-10-130
1161 Chu JSC, Tarailo-Graovac M, Zhang D, Wang J, Uyar B, Tu D, Trinh J, Baillie
1162 DL, Chen N. 2012. Fine tuning of RFX/DAF-19-regulated target gene
1163 expression through binding to multiple sites in Caenorhabditis elegans.
1164 *Nucleic Acids Res* **40**:53–64. doi:10.1093/nar/gkr690
1165 Chung MI, Peyrot SM, LeBoeuf S, Park TJ, McGary KL, Marcotte EM,
1166 Wallingford JB. 2012. RFX2 is broadly required for ciliogenesis during
1167 vertebrate development. *Dev Biol* **363**:155–165.
1168 doi:10.1016/j.ydbio.2011.12.029
1169 Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, Hilton JA,
1170 Jain K, Baymuradov UK, Narayanan AK, Onate KC, Graham K, Miyasato

SR, Dreszer TR, Strattan JS, Jolanki O, Tanaka FY, Cherry JM. 2018. The Encyclopedia of DNA elements (ENCODE): Data portal update. *Nucleic Acids Res*. Jan 4;46(D1):D794-D801. doi:10.1093/nar/gkx1081

De Stasio EA, Mueller KP, Bauer RJ, Hurlburt AJ, Bice SA, Scholtz SL, Phirke P, Sugiaman-trapman D, Stinson LA, Swoboda P. 2018. An Expanded Role for the RFX Transcription Factor. *Genetics* **208**:1083–1097. doi:10.1534/genetics.117.300571/-/DC1.1

Defrance M, Janky R, Sand O, van Helden J. 2008. Using RSAT oligo-analysis and dyad-analysis tools to discover regulatory signals in nucleic sequences. *Nat Protoc*. 3(10):1589-603. doi:10.1038/nprot.2008.98

Didon L, Zwick RK, Chao IW, Walters MS, Wang R, Hackett NR, Crystal RG. 2013. RFX3 Modulation of FOXJ1 regulation of cilia genes in the human airway epithelium. *Respir Res* **14**:1–13. doi:10.1186/1465-9921-14-70

Dubruille R, Laurençon A, Vandaele C, Shishido E, Coulon-Bublex M, Swoboda P, Couble P, Kernan M, Durand B. 2002. Drosophila regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* **129**:5487–5498. doi:10.1242/dev.00148

Edlund RK, Birol O, Groves AK. 2015. The Role of Foxi Family Transcription Factors in the Development of the Ear and Jaw. Current Topics in Developmental Biology. 111:461-95. doi:10.1016/bs.ctdb.2014.11.014

Efimenko E, Bubb K, Mark HY, Holzman T, Leroux MR, Ruvkun G, Thomas JH, Swoboda P. 2005. Analysis of xbx genes in C. elegans. *Development* **132**:1923–1934. doi:10.1242/dev.01775

Etchberger J, Lorch A, Sleumer M, Zapf R, Jones S, Marra M, Holt R, Moerman D, Hobert O. 2007. The molecular signature and cis-regulatory architecture of a *C. elegans* gustatory neuron. Genes & *Dev*. Jul 1;21(13):1653-74. doi:10.1101/gad.1560107.

Flames N, Hobert O. 2009. Gene regulatory logic of dopamine neuron differentiation. *Nature*. 458(7240):885-9. doi:10.1038/nature07929

Frøkjær-Jensen C, Ailion M, Lockery SR. 2008. Ammonium-acetate is sensed by gustatory and olfactory neurons in Caenorhabditis elegans. *PLoS One* **3** (6):e2467. doi:10.1371/journal.pone.0002467

Fujiwara M, Ishihara T, Katsura I. 1999. A novel WD40 protein, CHE-2, acts cell-autonomously in the formation of *C* . elegans sensory cilia. *Development* **126**:4839–4848. doi:10.1242/dev.126.21.4839

Garg A, Futcher B, Leatherwood J. 2015. A new transcription factor for mitosis: In Schizosaccharomyces pombe, the RFX transcription factor Sak1 works with forkhead factors to regulate mitotic expression. *Nucleic Acids Res* **43**:6874–6888. doi:10.1093/nar/gkv274

GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site—NDRI, Biospecimen Collection Source Site—RPCI, Biospecimen Core Resource—VARI, Brain Bank Repository—University of Miami Brain Endowment Bank, Leidos Biomedical—Project Management, ELSI Study, Genome Browser Data Integration &Visualization—EBI, Genome Browser Data Integration &Visualization—UCSC Genomics Institute, University of California Santa Cruz, Lead analysts:, Laboratory, Data Analysis &Coordinating Center (LDACC):, NIH program management:, Biospecimen

collection:, Pathology:, eQTL manuscript working group:, Battle A, Brown CD, Engelhardt BE, Montgomery SB. 2017. Genetic effects on gene expression across human tissues. *Nature* **550**:204–213. doi:10.1038/nature24277

Guo M, Wu T-H, Song Y-X, Ge M-H, Su C-M, Niu W-P, Li L-L, Xu Z-J, Ge C-L, Al-Mhanawi MTH, Wu S-P, Wu Z-X. 2015. Reciprocal inhibition between sensory ASH and ASI neurons modulates nociception and avoidance in Caenorhabditis elegans. *Nat Commun* **6**:5655. doi:10.1038/ncomms6655

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. 2007. Quantifying similarity between motifs. *Genome Biol* **8**. doi:10.1186/gb-2007-8-2-r24

Haycraft CJ, Swoboda P, Taulman PD, Thomas JH, Yoder BK. 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**:1493–505. doi:10.1242/dev.128.9.1493

Hilliard MA, Bargmann CI, Bazzicalupo P. 2002. C. elegans responds to chemical repellents by integrating sensory inputs from the head and the tail. *Current Biology* **12**:730–734. doi:10.1016/S0960-9822(02)00813-8

Hilliard MA, Bergamasco C, Arbucci S, Plasterk RHA, Bazzicalupo P. 2004. Worms taste bitter: ASH neurons, QUI-1, GPA-3 and ODR-3 mediate quinine avoidance in Caenorhabditis elegans. *EMBO Journal* **23**:1101–1111. doi:10.1038/sj.emboj.7600107

Hobert O. 2002. PCR fusion-based approach to create reporter Gene constructs for expression analysis in transgenic C. elegans. *Biotechniques* **32**:728–730. doi:10.2144/02324bm01

Hobert O, Moerman DG, Clark KA, Beckerle MC, Ruvkun G. 1999. A conserved LIM protein that affects muscular adherens junction integrity and mechanosensory function in Caenorhabditis elegans. *Journal of Cell Biology* **144**:45–57. doi:10.1083/jcb.144.1.45

Horani A, Ferkol TW. 2021. Understanding Primary Ciliary Dyskinesia and Other Ciliopathies. *Journal of Pediatrics* **230**:15-22.e1. doi:10.1016/j.jpeds.2020.11.040

Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MaCdonald J, Obenchain V, Oleš AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. 2015. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 12(2):115-21. doi:10.1038/nmeth.3252

Hutter H, Suh J. 2016. GExplore 1.4: An expanded web interface for queries on Caenorhabditis elegans protein and gene function . *Worm* **5(4)**:e1234659. doi:10.1080/21624054.2016.1234659

Kalderon D, Roberts BL, Richardson WD, Smith AE. 1984. A short amino acid sequence able to specify nuclear location. *Cell* **39**:499–509. doi:10.1016/0092-8674(84)90457-4

Kaplan JM, Horvitz HR. 1993. A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **90**:2227–2231. doi:10.1073/pnas.90.6.2227

Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, Van Der Lee R, Bessy A, Chèneby J, Kulkarni SR, Tan G, Baranasic D, Arenillas DJ, Sandelin A, Vandepoele K, Lenhard B, Ballester B, Wasserman WW, Parcy F, Mathelier A. 2018. JASPAR 2018: Update of the open-access

database of transcription factor binding profiles and its web framework. *Nucleic Acids Res*. 46(D1):D260-D266. doi:10.1093/nar/gkx1126
Kim H, Ishidate T, Ghanta KS, Seth M, Conte D, Shirayama M, Mello CC. 2014. A Co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. *Genetics* **197**:1069–1080. doi:10.1534/genetics.114.166389
Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, Almeida-King J, Staines D, Derwent P, Kerhornou A, Kersey P, Flicek P. 2011. Ensembl BioMarts: A hub for data retrieval across taxonomic space. *Database*. bar030. doi:10.1093/database/bar030
Kuleshov M V., Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 44(W1):W90-7. doi:10.1093/nar/gkw377
Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, Degnan BM. 2008. Genesis and expansion of metazoan transcription factor gene classes. *Mol Biol Evol* **25**:980–996. doi:10.1093/molbev/msn047
Lewis M, Stracker TH. 2020. Transcriptional regulation of multiciliated cell differentiation. *Semin Cell Dev Biol* **110**:51–60. doi:10.1016/j.semcdb.2020.04.007
Li JB, Gerdes JM, Haycraft CJ, Fan Y, Teslovich TM, May-Simera H, Li H, Blacque OE, Li L, Leitch CC, Lewis RA, Green JS, Parfrey PS, Leroux MR, Davidson WS, Beales PL, Guay-Woodford LM, Yoder BK, Stormo GD, Katsanis N, Dutcher SK. 2004. Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. *Cell* **117**:541–552. doi:10.1016/S0092-8674(04)00450-7
Li W, Kang L, Piggott BJ, Feng Z, Xu XZS. 2011. The neural circuits and sensory channels mediating harsh touch sensation in Caenorhabditis elegans. *Nat Commun* **2**:315. doi:10.1038/ncomms1308
Liu Y, Pathak N, Kramer-Zucker A, Drummond IA. 2007. Notch signaling controls the differentiation of transporting epithelia and multiciliated cells in the zebrafish pronephros. *Development* **134**:1111–1122. doi:10.1242/dev.02806
Lucas JS, Davis SD, Omran H, Shoemark A. 2020. Primary ciliary dyskinesia in the genomics age. *Lancet Respir Med* **8**:202–216. doi:10.1016/S2213-2600(19)30374-1
Luo Y, Hitz BC, Gabdank I, Hilton JA, Kagda MS, Lam B, Myers Z, Sud P, Jou J, Lin K, Baymuradov UK, Graham K, Litton C, Miyasato SR, Strattan JS, Jolanki O, Lee J-W, Tanaka FY, Adenekan P, O'Neill E, Cherry JM. 2020. New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res* **48**:D882–D889. doi:10.1093/nar/gkz1062
Margie O, Palmer C, Chin-Sang I. 2013. C. elegans chemotaxis assay. *Journal of Visualized Experiments*. Apr 27;(74):e50069. doi:10.3791/50069
Mello CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* **10**:3959–3970. doi:10.1002/j.1460-2075.1991.tb04966.x
Mitchell DR. 2017. Evolution of Cilia. *Cold Spring Harb Perspect Biol* **9**:a028290. doi:10.1101/cshperspect.a028290

Mukhopadhyay A, Deplancke B, Walhout AJM, Tissenbaum HA. 2005. C. elegans tubby regulates life span and fat storage by two independent mechanisms. *Cell Metab* **2**:35–42. doi:10.1016/j.cmet.2005.06.004

Mukhopadhyay S, Lu Y, Qin H, Lanjuin A, Shaham S, Sengupta P. 2007. Distinct IFT mechanisms contribute to the generation of ciliary structural diversity in C. elegans. *EMBO Journal* **26**:2966–2980. doi:10.1038/sj.emboj.7601717

Nakagawa S, Gisselbrecht SS, Rogers JM, Hartl DL, Bulyk ML. 2013. DNA-binding specificity changes in the evolution of forkhead transcription factors. *Proc Natl Acad Sci U S A* **110**:12349–12354. doi:10.1073/pnas.1310430110

Narasimhan K, Lambert SA, Yang AW, Riddell J, Mnaimneh S, Zheng H, Albu M, Najafabadi HS, Reece-Hoyes JS, Fuxman Bass JI, Walhout AJ, Weirauch MT, Hughes TR. 2015. Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities. *Elife* **4**. Apr 23;4:e06967. doi:10.7554/eLife.06967

Niwa S. 2017. Immobilization of Caenorhabditis elegans to Analyze Intracellular Transport in Neurons. *Journal of Visualized Experiments*. Oct 18;(128):56690. doi:10.3791/56690

Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, Waterston RH, Murray JI. 2019. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution . *Science Sep 20;**365**(6459):eaax1971*. doi:10.1126/science.aax1971

Perkins LA, Hedgecock EM, Thomson JN, Culotti JG. 1986. Mutant sensory cilia in the nematode Caenorhabditis elegans. *Dev Biol* **117**:456–87. doi:10.1016/0012-1606(86)90314-3

Phirke P, Efimenko E, Mohan S, Burghoorn J, Crona F, Bakhoum MW, Trieb M, Schuske K, Jorgensen EM, Piasecki BP, Leroux MR, Swoboda P. 2011. Transcriptional profiling of C. elegans DAF-19 uncovers a ciliary base-associated protein and a CDK/CCRK/LF2p-related kinase required for intraflagellar transport. *Dev Biol* **357**:235–247. doi:10.1016/j.ydbio.2011.06.028

Piasecki BP, Burghoorn J, Swoboda P. 2010. Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals. *Proceedings of the National Academy of Sciences* **107**:12969–12974. doi:10.1073/pnas.0914241107

Pierrou S, Hellqvist M, Samuelsson L, Enerbäck S, Carlsson P. 1994. Cloning and characterization of seven human forkhead proteins: Binding site specificity and DNA bending. *EMBO Journal* **13**:5002–5012. doi:10.1002/j.1460-2075.1994.tb06827.x

Quigley IK, Kintner C. 2017. Rfx2 Stabilizes Foxj1 Binding at Chromatin Loops to Enable Multiciliated Cell Gene Expression. *PLoS Genet* **13**:1–29. doi:10.1371/journal.pgen.1006538

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol*. Jan;29(1):24-6. doi:10.1038/nbt.1754

Sambongi Y, Nagae T, Liu Y, Yoshimizu T, Takeda K, Wada Y, Futai M. 1999. Sensing of cadmium and copper ions by externally exposed ADL, ASE, and

ASH neurons elicits avoidance response in Caenorhabditis elegans. *Neuroreport* **10**:753–757. doi:10.1097/00001756-199903170-00017

Sawin E, Ranganathan R, Horvitz H. 2000. C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. *Neuron*. Jun;26(3):619-31. doi: 10.1016/s0896-6273(00)81199-x.

Schiebinger G, Shu J, Tabaka M, Cleary B, Subramanian V, Solomon A, Gould J, Liu S, Lin S, Berube P, Lee L, Chen J, Brumbaugh J, Rigollet P, Hochedlinger K, Jaenisch R, Regev A, Lander ES. 2019. Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell* **176**:928-943.e22. doi:10.1016/j.cell.2019.01.006

Schindelin J, Arg I, Arganda-Carreras I, a-Carreras, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White D, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. Jun 28;9(7):676-82. doi: 10.1038/nmeth.2019.

Scholey J. 2007. The sensory cilia of Caenorhabditis elegans_Revised. *WormBook* **8**:2903–2915. doi:10.1895/wormbook.1.126.2

Sengupta P, Chou JH, Bargmann CI. 1996. odr-10 Encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. *Cell* **84**:899–909. doi:10.1016/S0092-8674(00)81068-5

Senti G, Swoboda P. 2008. Distinct Isoforms of the RFX Transcription Factor DAF-19 Regulate Ciliogenesis and Maintenance of Synaptic Activity. *Mol Biol Cell* **19**:5517–5528. doi:10.1091/mbc.e08-04-0416

Shapiro SS, Martin BW. 1965. An Analysis of Variance Test for Normality. *Biometrika*. Vol. 52, No. 3/4, pp. 591-611.

Shimeld SM, Degnan B, Luke GN. 2010. Evolutionary genomics of the Fox genes: Origin of gene families and the ancestry of gene clusters. *Genomics* **95**:256–260. doi:10.1016/j.ygeno.2009.08.002

Sorrells TR, Johnson AN, Howard CJ, Britton CS, Fowler KR, Feigerle JT, Weil PA, Johnson AD. 2018. Intrinsic cooperativity potentiates parallel cis-regulatory evolution. *Elife* **7**:1–29. doi:10.7554/eLife.37563

Sousa E, Flames N. 2022. Transcriptional regulation of neuronal identity. *European Journal of Neuroscience* **55**:645–660. doi:10.1111/ejn.15551

Starich TA, Herman RK, Kari CK, Yeh WH, Schackwitz WS, Schuyler MW, Collet J, Thomas JH, Riddle DL. 1995. Mutations affecting in chemosensory neurons of Caenorhabditis elegans. *Genetics* **139**:171–188. doi:10.1093/genetics/139.1.171

Stubbs JL, Oishi I, Izpisúa Belmonte JC, Kintner C. 2008. The forkhead protein Foxj1 specifies node-like cilia in Xenopus and zebrafish embryos. *Nat Genet* **40**:1454–1460. doi:10.1038/ng.267

Sugiaman-Trapman D, Vitezic M, Jouhilahti E-M, Mathelier A, Lauter G, Misra S, Daub CO, Kere J, Swoboda P. 2018. Characterization of the human RFX transcription factor family by regulatory and target gene analysis. *BMC Genomics* **19**:181. doi:10.1186/s12864-018-4564-6

Swoboda P, Adler HT, Thomas JH. 2000. The RFX-Type Transcription Factor DAF-19 Regulates Sensory Neuron Cilium Formation in C. elegans. *Mol Cell* **5**:411–421. doi:10.1016/S1097-2765(00)80436-0

Tarashansky AJ, Musser JM, Khariton M, Li P, Arendt D, Quake SR, Wang B. 2021. Mapping single-cell atlases throughout metazoa unravels cell type evolution. *Elife* **10**:1–24. doi:10.7554/eLife.66747

Taylor SR, Santpere G, Weinreb A, Barrett A, Reilly MB, Xu C, Varol E, Oikonomou P, Glenwinkel L, McWhirter R, Poff A, Basavaraju M, Rafi I, Yemini E, Cook SJ, Abrams A, Vidal B, Cros C, Tavazoie S, Sestan N, Hammarlund M, Hobert O, Miller DM. 2021. Molecular topography of an entire nervous system. *Cell* **184**:4329-4347.e23. doi:10.1016/j.cell.2021.06.023

Thevenon J, Duplomb L, Phadke S, Eguether T, Saunier A, Avila M, Carmignac V, Bruel AL, St-Onge J, Duffourd Y, Pazour GJ, Franco B, Attie-Bitach T, Masurel-Paulet A, Rivière JB, Cormier-Daire V, Philippe C, Faivre L, Thauvin-Robinet C. 2016. Autosomal recessive IFT57 hypomorphic mutation cause ciliary transport defect in unclassified oral–facial–digital syndrome with short stature and brachymesophalangia. *Clin Genet* **90**:509–517. doi:10.1111/cge.12785

Thomas J, Morlé L, Soulavie F, Laurençon A, Sagnol S, Durand B. 2010. Transcriptional control of genes involved in ciliogenesis: a first step in making cilia. *Biol Cell* **102**:499–513. doi:10.1042/bc20100035

Thomas-Chollier M, Darbo E, Herrmann C, Defrance M, Thieffry D, Van Helden J. 2012a. A complete workflow for the analysis of full-size ChIP-seq (and similar) data sets using peak-motifs. *Nat Protoc*. 7(8):1551-68. doi:10.1038/nprot.2012.088

Thomas-Chollier M, Herrmann C, Defrance M, Sand O, Thieffry D, Van Helden J. 2012b. RSAT peak-motifs: Motif analysis in full-size ChIP-seq datasets. *Nucleic Acids Res*. Feb;40(4):e31. doi:10.1093/nar/gkr1104

Timpson NJ, Greenwood CMT, Soranzo N, Lawson DJ, Richards JB. 2018. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat Rev Genet* **19**:110–124. doi:10.1038/nrg.2017.101

Tobin JL, Beales PL. 2009. The nonmotile ciliopathies. *Genetics in Medicine* **11**:386–402. doi:10.1097/GIM.0b013e3181a02882

Troemel ER, Kimmel BE, Bargmann CI. 1997. Reprogramming chemotaxis responses: Sensory neurons define olfactory preferences in C. elegans. *Cell* **91**:161–169. doi:10.1016/S0092-8674(00)80399-2

Turatsinze JV, Thomas-Chollier M, Defrance M, van Helden J. 2008. Using RSAT to scan genome sequences for transcription factor binding sites and cis-regulatory modules. *Nat Protoc*. 3(10):1578-88. doi:10.1038/nprot.2008.97

Vij S, Rink JC, Ho HK, Babu D, Eitel M, Narasimhan V, Tiku V, Westbrook J, Schierwater B, Roy S. 2012. Evolutionarily Ancient Association of the FoxJ1 Transcription Factor with the Motile Ciliogenic Program. *PLoS Genet* **8**:e1003019. doi:10.1371/journal.pgen.1003019

Wagner GP. 2015. Two rules for the detection and quantification of epistasis and other interaction effects. *Methods in Molecular Biology*. 1253:145-57. doi: 10.1007/978-1-4939-2155-3_8.

Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, Zheng H, Goity A, van Bakel H, Lozano JC, Galli M, Lewsey MG, Huang E, Mukherjee T, Chen X, Reece-Hoyes JS, Govindarajan S, Shaulsky G, Walhout AJM, Bouget FY, Ratsch G, Larrondo LF, Ecker JR, Hughes TR. 2014. Determination and

inference of eukaryotic transcription factor sequence specificity. *Cell* **158**:1431–1443. doi:10.1016/j.cell.2014.08.009

Yu G, Wang LG, He QY. 2015. ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. 31(14):2382-3. doi:10.1093/bioinformatics/btv145

Yu X, Ng CP, Habacher H, Roy S. 2008. Foxj1 transcription factors are master regulators of the motile ciliogenic program. *Nat Genet* **40**:1445–1453. doi:10.1038/ng.263

Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, Häring M, Braun E, Borm LE, La Manno G, Codeluppi S, Furlan A, Lee K, Skene N, Harris KD, Hjerling-Leffler J, Arenas E, Ernfors P, Marklund U, Linnarsson S. 2018. Molecular Architecture of the Mouse Nervous System. *Cell* **174**:999-1014.e22. doi:10.1016/j.cell.2018.06.021

Zhang C, Zhao N, Chen Y, Zhang D, Yan J, Zou W, Zhang K, Huang X. 2016. The signaling pathway of Caenorhabditis elegans mediates chemotaxis response to the attractant 2-heptanone in a Trojan Horse-like pathogenesis. *Journal of Biological Chemistry* **291**:23618–23627. doi:10.1074/jbc.M116.741132

Zhang F, Bhattacharya A, Nelson JC, Abe N, Gordon P, Lloret-Fernandez C, Maicas M, Flames N, Mann RS, Colón-Ramos DA, Hobert O. 2014. The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types. *Development* **141**:422–435. doi:10.1242/dev.099721

FIGURE LEGENDS

Figure 1. FKH-8 is a candidate for direct regulation of ciliome gene expression in *C. elegans.*

A) Schema for a sensory cilium. Cilia components (ciliome) can be divided into core and subtype specific categories. Core genes whose reporters are analysed in panel C and Figure 1 Figure Supplement 2 are indicated by their function.

B) Left lateral view of *C. elegans* hermaphrodite ciliated system. Sixty ciliated neurons from 25 different classes are distributed in 5 distinct anatomical regions.

C) Ciliome core components show persistent expression in double *daf-12(sa204); daf-19(m86)* mutants. The same extrachromosomal line was analysed in the different genetic backgrounds. Each dot represents the total number of reporter-positive neurons in a single animal. Mean and standard error are represented. The mean number of remaining reporter-positive neurons in double *daf-12; daf-19* mutants is indicated. Sample sizes for each genetic background: *che-11: n= 8; osm-1: n ≥7; ift-20: n ≥9; che-13: n ≥7; tub-1: n ≥10; xbx-1: n ≥5; peli-1: n ≥8.* See **Figure 1 Source Data 1** for raw data and **Figure**

1 Figure Supplement 1 and 2 for details on construct lengths and additional reporter scorings.

D) sc-RNA-seq data analysis identifies 10 TFs specifically enriched in ciliated sensory neurons. These TFs belong to FKH, ZF, NHR and HD families. See **Figure 1 Figure Supplement 3** for detailed description of TF expression in each ciliated neuron type.

E) ChIP-seq data analysis of 259 available TFs shows that FKH-8 ranks first in direct binding to regulatory regions assigned to the ciliome gene list. See **Figure 1 Source Data 2** for gene lists and **Figure 1 Figure Supplement 3** for core ciliome or subtype specific binding analysis.

F) Correlation of total number of peaks versus ciliome-gene peaks shows FKH-8 behaves as an outlier, demonstrating high binding to ciliome genes is not merely due to the high number of FKH-8 binding-events.

Figure 2. FKH-8 is expressed in sensory ciliated neurons, binds ciliome genes near DAF-19 X-boxes and physically interacts with DAF-19.

A) *fkh-8* locus (top) and fosmid based *fkh-8* reporter (bottom). Grey boxes represent exons and orange boxes correspond to exons coding for the FKH DNA binding domain (DBD). Putative *daf-19*/RFX binding sites (X-boxes) are depicted with blue lines. Red bars indicate extension for the corresponding deletion alleles.

B) Dorso-ventral views of young adult animals expressing both the fosmid-based FKH-8::GFP reporter (in green) and an integrated reporter for the panciliary marker *ift-20* (in red). A: anterior, P: posterior, R: right, L: left. Scale bar = 10 µm. See **Figure 2 Source Data 1** for quantification and **Figure 2 Figure Supplement 1** for embryonic expression patterns and expression correlation with DAF-19 and ciliome genes.

C) Genes associated to nearby FKH-8 binding events enrich Gene Ontology terms related to cilia regulated processes and/or functions. Data correspond to adjusted p-value. See **Figure 2 Source Data 2** for gene lists associated to GO terms

D) *De novo* motif analysis of FKH-8 ChIP-seq data identifies a motif present in 27% of peaks, enriched at central positions, that matches a Forkhead like (FHL) motif.

E) DAF-19/RFX binding motifs (PWM M1534_1.02) are present in 21% of the FKH-8 bound regions and are enriched at central positions. See **Figure 2 Figure Supplement 2** for similar analysis on additional FKH ChIP-seq data sets.
F) Distribution of ciliome genes in four different categories: 1) genes with both X-box motifs and FKH-8 binding events; 2) genes with only FKH-8 binding; 3) Genes with X-box motifs only and 4) Genes with neither FKH-8 binding or X-boxes. Most ciliome genes contain both X-boxes and FKH-8 peaks, this dual signature is highly prevalent in core ciliome genes while is minoritary in subtype ciliome genes. See **Figure 2 Source Data 2** for gene lists associated to each signature.
G) Distance between X-boxes found in ciliome genes and the center of the nearest FKH-8 ChIP-seq peak. 42% of X-boxes are located less than 600 bp from a FKH-8 ChIP-seq peak. See **Figure 2 Figure Supplement 2** for differential analysis of core and subtype ciliome genes.
H) Co-immuno precipitation of HA tag FKH-8 and FLAG tag DAF-19 expressed in HEK293 cells shows physical interaction between both transcription factors in the soluble fraction of nuclear extracts. MEK2 is used to assess for the presence of cytoplasmic components and Histone H3 to assess the presence of chromatin. See **Figure2 Source Data 3** for original blots and **Figure 2 Figure Supplement 2** for additional analysis of interaction in chromatin associated fractions.

Figure 3. FKH-8 TF and FKH-binding sites are required for correct core ciliome gene reporter expression.
A) Dorso-ventral images from young adult heads expressing different core ciliome multicopy array gene reporters in *wild type* and *fkh-8(vlc43)* null mutant animals. All reporters are extrachromosomal arrays except for *ift-20* reporter which is integrated. Arrow heads point deirid expression lost in the mutant. A: anterior, P: posterior, R: right, L: left. Scale bar = 25 µm.
B) Quantification of the number of *gfp* positive cells in five distinct anatomical regions for each reporter in *wild type, fkh-8(tm292)* hypomorphic allele and *fkh-8(vlc43)* null mutant. To facilitate comparisons, values in each region are normalized to controls. The same extrachromosomal line was analysed in the different genetic backgrounds. Each dot represents the number of reporter-expressing neurons scored in a single animal. Mean and standard error are

represented. Black asterisk denotes significantly different from *wild type* and orange asterisk indicates *vlc43* is significantly different from *tm292* allele. Sample sizes for each genetic background: *ift-20: n≥10; osm-9: n=9; osm-5: n≥8; peli-1: n≥7; xbx-1: n≥5.* See **Figure 3 Source Data 1** for raw scoring data, **Figure 3 Figure Supplement 1** for analysis of the hypomorphic recessive nature of the *tm292* allele and quantification of additional reporters not affected in *fkh-8* mutants, see **Figure 3 Figure Supplement 2 and Figure 3 Source Data 2** for functional characterization of predicted FKH binding sites in *ift-20* and *xbx-1* regulatory regions.

C) Dorso-ventral images from young adult heads expressing *GFP* from the endogenously tagged *osm-5* locus *[osm-5(syb6528), osm-5::SL2::GFP::H2B]* in *wild type* and *fkh-8(vlc43)* null mutant. A global decrease in fluorescence intensity is detected in *fkh-8(vlc43)* animals compared to *wild type.*

D) Fluorescence intensity level quantification in specific ciliated neuron populations shows significant reduction of expression in *fkh-8(vlc43)* animals. A. U.: arbitrary units. See **Figure 3 Source Data 1** for raw scoring data. *n≥20* for each cell type and genetic background.

Figure 4. FKH-8 and DAF-19 exhibit crosstalk and synergistic effects in the transcriptional regulation of ciliome genes.

A) *daf-19* locus codes for five different *daf-19* isoforms. Grey boxes represent exons whereas blue boxes correspond to exons coding for the RFX DNA binding domain (DBD). FKH-8 binding events are depicted as orange lines. Red arrows locate mutations of the corresponding alleles.

B) Lateral views from young adult hermaphrodite heads expressing *fkh-8* fosmid-based reporter (*wgIs652)*. Lack of all *daf-19* isoforms (*m86* allele) derepresses *fkh-8* in non-ciliated neurons. This phenotype is mimicked by the specific absence of long *daf-19a/b* isoforms (*of5* allele). Scale bar = 25 µm. See **Figure 4 Figure Supplement 1** for unaffected DAF-19 expression in ciliated neurons in *fkh-8(vlc43)* mutants.

C) Dorso-ventral images from young adult hermaphrodites showing core ciliome *ift-20* reporter expression in different genetic backgrounds. Scale bar = 25 µm.

D) Mean number of *ift-20* reporter-expressing neurons in *daf-12(sa204); daf-19(m86)*, *fkh-8(tm292)* triple mutants is significantly different from each of the double mutants and significantly lower than the expected from the multiplicative effect of both *daf-12(sa204); fkh-8(tm292)* and *daf-12(sa204); daf-19(m86)* animals. The same extrachromosomal line was analysed in the different genetic backgrounds. Each dot represents the number of reporter-expressing neurons scored in a single animal. Mean and standard error are represented. See **Figure 4 Figure Supplement 2** for quantification of FKH-8 and DAF-19 synergistic effects in *xbx-1* and *peli-1* reporter expression and **Figure 4 Source Data 1** for raw data and statistics for all analyzed genetic backgrounds. $n \geq 10$ for each genetic background.

E) Analysis of *osm-1* and *xbx-1* ciliome reporters in specific subpopulations of ciliated neurons. CEPV and CEPD are labelled with *dat-1::mcherry (otIs181) and srh-211:: tagRFP* is expressed in PHA neuron, both reporters are unaffected in all genetic backgrounds. Quantification of ciliome reporters is depicted in the corresponding graphs. **Figure 4 Source Data 1** for raw data and statistics for all analyzed genetic backgrounds n=30 worms per genotype and reporter construct.

Figure 5. *fkh-8(vlc43)* null mutants display morphological defects in cilia.

Integrated reporters unaffected in *fkh-8* mutant are used to label the cilia of several distinct subpopulations of ciliated neurons. CEP: *otIs259(dat-1::gfp)*; ADF: *zdIs13(tph-1::gfp)*; AWB: *kyIs104(str-1::gfp)*; AWA: *pkIs583(gpa-6::gfp)*. Panels show representative images from 3 animals in *wild type* and *fkh-8(vlc43)* mutant backgrounds. Cilium length of CEP and AWB neurons is significantly reduced in the absence of FKH-8 whereas ADF cilia length is increased. Depth of AWA cilium arborization is significantly reduced in *fkh-8(vlc43)* null mutants. Each dot in the graphs represents measures for a single cilium. Mean and standard error are represented. See **Figure 5 Figure Supplement 1** for cilia morphology analysis in worms immobilized in polystyrene beads and **Figure 5 Source Data 1** for raw data and statistics. CEP and AWB n=40; ADF n=32; AWA n=7.

Figure 6. FKH-8 is required for the correct display of several sensory mediated behaviours.

A) Mutations in *fkh-8* significantly impair appropriate backward response to nose touch, revealing functionality defects for the ASH, FLP and/or OLQ ciliated neurons. This phenotype is stronger in *fkh-8(vlc43)* null mutants than in the hypomorphic *tm292* allele. n=20 animals per replicate, 3 biological replicates per genotype.

B) Decrease in locomotory rate upon re-entering a bacterial lawn is unaffected in *fkh-8* mutants. n=15 worms per genotype and condition.

C) *fkh-8* null mutants significantly fail to prevent *dauer* entry. Pheromones induce *dauer* in *fkh-8* mutants, albeit less efficiently than in controls. Four biological replicates n>295 per replicate and genotype.

D to F) Lack of *fkh-8* significantly impairs olfaction-mediated behaviours. Defects are observed for 2-nonanone repulsion mediated by AWB [*Wild type* n= 59, 128, 114, 165; *fkh-8(tm292)* n=76, 123, 129, 209 and *fkh-8(vlc43)* n=82, 92, 130, 139] and 2-heptanone attraction mediated by AWC neurons [*Wild type* n=124, 129, 133; *fkh-8(tm292)* n=68, 94, 102 and *fkh-8(vlc43)* n=87, 83, 85]. Diacetyl response, mediated by AWA, is affected but not to a significant level due to high variability in the response [*Wild type* n=168, 69, 103; *fkh-8(tm292)* n=57, 85, 110 and *fkh-8(vlc43)* n=115, 107, 74].

G to I) Attractive chemotaxis towards NaCl is unaffected in *fkh-8* mutant animals. [*Wild type* n=62, 78, 72; *fkh-8(tm292)* n=105, 116, 106 and *fkh-8(vlc43)* n=111, 52, 78]. Avoidance behaviour towards toxic SDS and copper anions is significantly impaired. [Six biological replicates, 5 worms per replicate and genotype, 4 tests per worm].

Mean and standard error are represented in all graphs. See **Figure 6 Figure Supplement 1** for quantification of non-cilia mediated behaviours and **Figure 6 Source Data 1** for raw data and statistics.

Figure 7. Mammalian FKH TFs with known motile cilia regulatory functions can rescue *fkh-8* mutant phenotype.

A) Rescue strategy: *dat-1* promoter, unaffected in *fkh-8* mutants, is used to drive FKH TF cDNA and eGFP expression specifically in the dopaminergic ciliated system.

B) Representative images of dopaminergic neurons expressing an integrated reporter for the core ciliome marker *ift-20* (in red) in *wild type*, *fkh-8(vlc43)*

mutants and with the co-expression of different rescuing constructs. Scale bar = 5 µm.

C) Quantification of rescue experiments. *ift-20* reporter expression is lost from the dopaminergic neurons in *fkh-8(vlc43)* null mutants compared to *wild type* animals. Expression of FKH-8, FOXJ1 and FOXN4 but not FOXI1 is sufficient to recover *ift-20* expression in dopaminergic neurons. N = 30 animals per transgenic line. See **Figure 7 Source Data 1** for raw data and similar results obtained with two additional transgenic lines per construct.

D) Speculative model on the evolution of ciliome gene regulatory logic. FKH and RFX TFs could have an ancestral role in the direct coregulation of ciliome genes before its functional diversification into motile and primary cilia cell types. Different RFX and FKH TF members could have evolved to regulate ciliome genes in specific cell types in different organisms. Orange squares represent FKH TFs and blue circles RFX TFs, light blue bars represent ciliome enhancers.

FIGURE SUPPLEMENTS LEGENDS

Figure 1 Figure Supplement 1. Ciliome reporters used in this work.

Schematic representation of reporter constructs used in the manuscript. Selected core cilia components contain at least one experimentally validated X-box motif in their sequences (marked as a blue bar). For *che-11*, *che-13*, *osm5*, *ift-20*, *tub-1*, *mks-1* and *osm-1* see (Efimenko et al., 2005); for *peli-1* see (Chu et al., 2012), for *xbx-1* see (Schafer et al., 2003); for *tmem-107* see (Lambacher et al., 2016). Overlap between x-boxes and FKH-8 binding sites is found for *osm-5*, *ift-20*, *peli-1*, *xbx-1*, *tub1* and *osm-1*.

Figure 1 Figure Supplement 2. Reporter expression for some core ciliome genes is abolished in *daf-19(m86)* mutant.

Expression of short reporters for the core cilia components *tmem-107*, *mks-1* and *osm-5* is completely abolished in double *daf-12(sa204)*; *daf-19(m86)* double mutants. The same extrachromosomal line was analysed in the different genetic backgrounds. $n \geq 10$.

Each dot represents the total number of reporter-positive neurons scored in a single animal. Mean and standard error are represented. See **Figure 1 Source Data 1** f for raw scoring data.

Figure 1 Figure Supplement 3. Available -omics data identifies FKH-8 as a candidate transcriptional regulator of ciliome genes in *C. elegans*.

A) DAF-19/RFX motifs (X-box) are enriched in regulatory sequences of core and broadly expressed ciliome genes more prominently than in subtype specific features. Motif logo for the palindromic X-box motif.

B) *De novo* motif enrichment analysis of putative regulatory sequences of ciliome genes identifies a motif matching known binding site for the pro-neural bHLH TFs *lin-32* and *hlh-11*.

C) Two additional motifs enriched in regulatory sequences of ciliome genes containing X-box sites, show partial similarity to the DAF-19/RFX motif and to the ciliome-related motif C-BOX motif (Burghoorn et al., 2012).

D) sc-RNA-seq data of FACS-isolated neurons from L4 hermaphrodites (Taylor et al., 2021) show broad expression for *ceh-57*, *fkh-8*, *nhr-158* and *nhr-277* TFs across the whole ciliated system of *C. elegans*. Only *fkh-8* expression is detected in all ciliated neuron types. Red arrows indicate values lower than 10 TPM (transcripts per million).

E) ChIP-seq data analysis shows FKH-8 ranks first among 259 TFs directly binding to either core ciliome genes (left) or subtype-specific ciliary features (right).

Figure 2 Figure Supplement 1. *fkh-8* expression along ciliated system development.

A) Representative Z projections of developmental embryonic milestones until hatching (L1) shows FKH-8::GFP fosmid reporter expression matches in time and space panciliary reporter *ift-20::gfp* expression. Scale bar = 10 µm.

B) Representative Z projections of two-fold embryo and fist larval stage (L1) animals expressing both FKH-8::GFP fosmid reporter (in green) and an integrated reporter for the panciliary marker *ift-20:: tagRFP* (in red). Note that due to long maturation time of the tag-RFP reporter, *ift-20::tagRFP* expression is only

1743 detected from the 2 fold stage, while *ift-20::gfp* reporter in **(**A**)** is first detected at
1744 bean stage, similar to *fkh-8* expression. Scale bar = 10 µm.
1745 C) Single Z-plane from regions indicated in (B) show colocalization of FKH-
1746 8::GFP and RFP in the ciliated sensory neurons.
1747 D) Embryonic sc-RNA-data (Packer et al., 2019) from *C. elegans* ciliated neurons
1748 and their progenitor cells. Pseudo-time (left pannel) shows the maturation
1749 trajectory of ciliated neurons that coincides with increasing *ift-20*, *fkh-8* and *daf-*
1750 *19* expression.
1751 E) Correlation index of *fkh-8, daf-19* and *hlh-14* TF scRNAseq expression and
1752 four different gene categories (core ciliome, subtype ciliome, panneuronal or
1753 ubiquitous) in all ciliated lineages (Packer et al., 2019). *fkh-8* and *daf-19*
1754 expression shows high correlation index with core ciliome genes but not with
1755 other gene categories, while *hlh-14*, bHLH TF not involved in ciliogenesis shows
1756 low correlation values in all categories. See **Figure 2 Source Data 2** f for raw
1757 data.
1758
1759 **Figure 2 Figure Supplement 2. FKH-8 binds near X-BOX motifs**
1760 A) Analysis of DAF-19/RFX binding motifs (PWM M1534_1.02) in peaks for ChIP-
1761 seq datasets of other *C. elegans* FKH TFs present in ENCODE database. There
1762 is no significant enrichment for DAF-19 motif in FKH-2, FKH-3 and FKH-10
1763 datasets. Motif enrichment in FKH-4, FKH-6 and PHA-4 datasets is less
1764 significant and shows less defined enrichment in the centre of the peaks. p
1765 represents P Value associated to motif enrichment and % refers to the number
1766 of peaks with the motif present. EM: Embryo mixed stage; L1: Larval stage 1; L3:
1767 Larval stage 3, L4: Larval stage 4, LE: Late embryo. See **Figure 2 Source Data**
1768 **2** for detailed data.
1769 B) Distance between X-boxes found in the promoter regions of core ciliome genes
1770 and the center of the nearest FKH-8 ChIP peak. 63% of X-boxes are located less
1771 than 600 bp from a FKH-8 ChIP peak.
1772 C) Distance between X-boxes found in the promoter regions of subtype ciliome
1773 genes and the center of the nearest FKH-8 ChIP peak. 9% of X-boxes are located
1774 less than 600 bp from a FKH-8 ChIP peak.
1775 D) Co-immuno precipitation of FKH-8 and DAF-19 expressed in HEK293 cells.
1776 Micrographs shows nuclear localization of transfected HA::FKH-8 and

FLAG::DAF-19. In addition to the interaction detected in the soluble nuclear fraction (Figure 2), both factors also interact bound to DNA (chromatin fraction). See **Figure 2 Figure Supplement 3 Source data 1** for original blots.

Figure 3 Figure Supplement 1. *f. fkh-8(tm292)* is a hypomorphic recessive allele.

A) Quantification of the number of *gfp* positive cells in 5 distinct anatomical regions for each reporter in *wild type, fkh-8(tm292)* hypomorphic allele and and *fkh-8(vlc43)* null mutant. To facilitate comparisons, values in each region are normalized to controls. The same extrachromosomal line was analysed in the different genetic backgrounds. Each dot represents the number of reporter-expressing neurons scored in a single animal. Mean and standard error are represented. *fkh-8(tm292)* and *fkh-8(vlc43)* show similar expression values than wild type for these markers. See **Figure 3 Source data 1** for raw scoring data.

B) Quantification of the number of *ift-20::rfp* positive cells in *wild type*, *fkh-8(tm292)* and *fkh-8(vlc43)* mutants shows stronger defects in *vlc43* null allele. *: n ≥ 5.*

C) Heterozygote cross progeny from *wild type*, *fkh-8(tm292)* and *fkh-8(vlc43)* show similar number of *ift-20::rfp* positive cells denoting the recessive nature of both alleles. *n ≥ 10.*

Figure 3 Figure Supplement 2. *f.* Functional characterization of putative FKH sites in *cis*-regulatory modules of two core ciliome components.

A) Schematics for the *ift-20* and *xbx-1* loci and reporters. Dark grey boxes represent exons whereas light grey boxes correspond to UTRs. FKH-8 peaks are depicted with an orange horizontal line while predicted FKH DNA binding motifs are indicated with a vertical orange bar. Sequences corresponding to *wild type* and mutated putative FKH sites are indicated. See **Figure 3 Source data 2** for FKH putative binding site assignment.

B) *Cis*-regulatory mutation of putative FKH sites greatly reduces ciliome gene reporter expression. Representative dorso-ventral images from young adult heads expressing *wild type* or FKH-site-mutated reporters for core ciliome genes *ift-20 (*left) and *xbx-1* (right). A: anterior, P: posterior, R: right, L: left. Scale bar = 25 µm.

C**)** Quantification of total number of *gfp* positive cells for *wild type* and point mutated *gfp* reporters. Each graph represents one of the five anatomical regions scored. Three different extrachromosomal lines were analysed for each construct. Each dot represents the number of reporter-expressing neurons scored in a single animal. Statistically significant differences are indicated with asterisks. See **Figure 3 Source data 1** for raw scoring data. n=10.

Figure 4 Figure Supplement 1. Lack of FKH-8 has no major effect on DAF-19 expression.

A) Representative lateral views from heads of young adult hermaphrodites co-expressing a fosmid-based DAF-19::GFP reporter and *dat-1::mcherry* reporter labelling the dopaminergic neurons. Lack of FKH-8 does not seem to affect DAF-19::GFP expression pattern. Co-localization analysis shows normal expression in the dopaminergic ciliated neurons (CEPV, CEPD, ADE, PDE), quantified in the graphs. Scale bar = 20 μm. See **Figure 4 Source data 1** for raw data. n=50.

B) *daf-19* expression is largely unaffected in the subpopulation of DiD-positive ciliated amphid neurons in null *fkh-8* mutant animals. DAF-19::GFP is consistently detected in the ASI, ADL and AWB neurons in both *wild type* and null *fkh-8* mutant backgrounds. Mean and standard deviation are represented. n = 10 animals.

Figure 4 Figure Supplement 2. FKH-8 and DAF-19 show synergistic effects in the transcriptional regulation of the ciliome.

A) Representative micrographs of *kap-1* reporter expression in different genetic backgrounds. *kap-1::gfp* is unaffected in *fkh-8(tm292)* and ectopically expressed both in *daf-19(m86),daf-12(sa204)* and *daf-19(m86),daf-12(sa204), fkh-8(tm292).* Ectopic expression is more evident in the posterior part of the head, labelled with a white line.

B) Quantification of *kap-1* reporter expression in different subpopulations of sensory ciliated neurons shows similar expression in *fkh-8(vlc43)* null mutants and *wild type* animals. n ≥ 19.

C) Representative micrographs of embryonic *ift-20::rfp* expression after ectopic expression of DAF-19D and FKH-8 through a heatshock inducible promoter. 2 cell stage embryos were grown at 20ºC for 4 hours and then heat shocked at

37ºC for 20 minutes, embryos were scored for ectopic *ift-20::rfp* 20 hours after heatshock. Two different lines were analysed for each construct. 100% of embryos with DAF19D ectopic expression show ectopic *ift-20::rfp* expression, while heatshocked embryos without array had normal *ift-20::rfp* expression (Line 1 n=24 embryos with array, n=11 embryos without array; Line 2 n=21 embryos with array, n=12 embryos without array). Ectopic FKH-8 expression had no obvious effect on *ift-20::rfp.* (Line 1 n=10 embryos with array, n=5 embryos without array; Line 2 n=15 embryos with array, n=3 embryos without array).
D) Quantification of ciliome reporters in *fkh-8* and *daf-19* mutants. The same extrachromosomal line was analysed in the different genetic backgrounds. Each dot represents the number of reporter-expressing neurons scored in a single animal. Mean and standard error are represented. For both reporters triple mutant is significantly different from each of the double mutants and significantly lower than the expected from the multiplicative effect of both *daf-12(sa204)*; *fkh-8(tm292)* and *daf-12(sa204); daf-19(m86)* animals. *n ≥ 6.*

Figure 5 Figure Supplement 1. *fkh-8(vlc43)* null mutants display morphological defects in cilia using physical immobilization with polystyrene beads.
Integrated reporters unaffected in *fkh-8* mutant are used to label the cilia of several distinct subpopulations of ciliated neurons. CEP: *otIs259(dat-1::gfp)*; ADF: *zdIs13(tph-1::gfp)*; AWB: *kyIs104(str-1::gfp)*; AWA: *pkIs583(gpa-6::gfp)*. Panels show representative images from *wild type* and *fkh-8(vlc43)* mutant backgrounds. Cilium length of CEP ADF and PHB is significantly reduced in the absence of FKH-8 whereas AWB cilia length is unaffected. Mean and standard error are represented. Please not different fixation methods are used compared to Figure 5. *CEP, AWB: n ≥ 50, ADF: n ≥ 30, PHB: n ≥ 28.*

Figure 6 Figure Supplement 1. FKH-8 is not required for correct display of mechanosensory behaviours mediated by non-ciliated neurons.
A to C) *fkh-8* mutants show normal avoidance behaviours elicited by mechanical stimuli known as gentle touch and harsh touch paradigms, suggesting FKH-8 is not required for the correct functionality of non-ciliated neurons ALM, AVM, PLM and PVD. Redundant actions of PVD and PDE controlling scape response to

harsh touch prevent to assess defects about the functionality of ciliated PDE neurons. Mean and standard error for three independent replicates are represented. See **Figure 6 Source data 1** for raw data. n=60 worms per genotype.

Figure 1 Source Data 1. Raw quantification data of ciliome gene reporters in *daf-19(m86); daf-12(sa204)* included in Figure 1 and Figure 1 Figure Supplement 2

Figure 1 Source Data 2. Gene lists and raw data for bioinformatics analysis in Figure 1 and Figure 1 Figure Supplement 3.

Figure 2 Source Data 1. Raw quantification data of FKH-8 and ift-20 reporter co-expression represented in Figure 2B.

Figure 2 Source Data 2. Raw data for bioinformatics analysis in Figure 2 and Figure 2 Figure Supplement 1 and 2.

Figure 2 Source Data 3. Co-IP original files of unedited gels in Figure 2.

Figure 3 Source Data 1. Raw quantification data of ciliome gene reporter expression defects in *fkh-8* mutants and cis-regulatory point mutation analysis corresponding to Figure 3 and Figure 3 Figure Supplement 1 and 2

Figure 3 Source Data 2. Motif enrichment analysis of *xbx-1* and *ift-20* reporters.

Figure 4 Source Data 1. Raw quantification data of synergistic actions of fkh-8 and daf-19 depicted in Figure 4, Figure 4 Figure Supplement 1 and 2.

Figure 5 Source Data 1. Raw quantification data of cilia morphology analysis depicted in Figure 5 and Figure 5 Figure Supplement 1.

Figure 6 Source Data 1. Raw quantification data for behavioural analysis in Figure 6 and Figure 6 Figure Supplement 1.

Figure 7 Source Data 1. Raw quantification data for rescuing experiments in Figure 7.

Supplementary File 1. List of reagents: Strains, plasmids and primers.

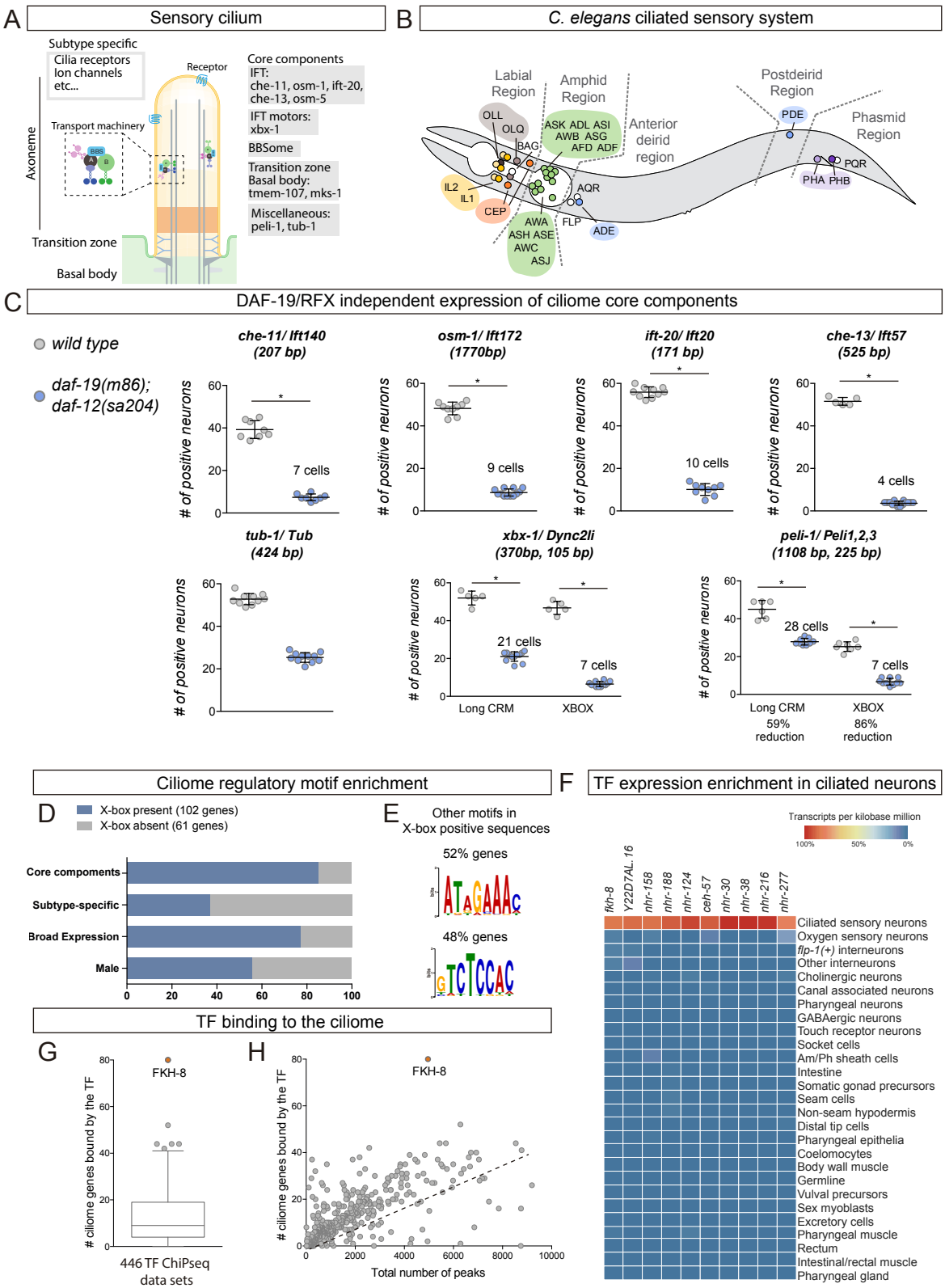
A
Sensory cilium
Subtype specific
Cilia receptors
Ion channels
etc...
Receptor
Transport machinery
Axoneme
Transition zone
Basal body
Core components
IFT:
che-11, osm-1, ift-20,
che-13, osm-5
IFT motors:
xbx-1
BBSome
Transition zone
Basal body:
tmem-107, mks-1
Miscellaneous:
peli-1, tub-1
B
C. elegans ciliated sensory system
Labial Region
Amphid Region
Anterior deirid region
Postdeirid Region
Phasmid Region
OLL
OLQ
BAG
ASK ADL ASI
AWB ASG
AFD ADF
PDE
PQR
PHA PHB
IL2
IL1
CEP
AQR
FLP
ADE
AWA
ASH ASE
AWC
ASJ
C
DAF-19/RFX independent expression of ciliome core components
wild type
daf-19(m86); daf-12(sa204)
che-11/ Ift140 (207 bp)
osm-1/ Ift172 (1770bp)
ift-20/ Ift20 (171 bp)
che-13/ Ift57 (525 bp)
# of positive neurons
7 cells
9 cells
10 cells
4 cells
tub-1/ Tub (424 bp)
xbx-1/ Dync2li (370bp, 105 bp)
peli-1/ Peli1,2,3 (1108 bp, 225 bp)
21 cells
7 cells
28 cells
7 cells
Long CRM
XBOX
59% reduction
86% reduction
Ciliome regulatory motif enrichment
D
X-box present (102 genes)
X-box absent (61 genes)
Core compoments
Subtype-specific
Broad Expression
Male
E
Other motifs in X-box positive sequences
52% genes
48% genes
F
TF expression enrichment in ciliated neurons
Transcripts per kilobase million
fkh-8
Y22D7AL.16
nhr-158
nhr-188
nhr-124
ceh-57
nhr-30
nhr-38
nhr-216
nhr-277
Ciliated sensory neurons
Oxygen sensory neurons
flp-1(+) interneurons
Other interneurons
Cholinergic neurons
Canal associated neurons
Pharyngeal neurons
GABAergic neurons
Touch receptor neurons
Socket cells
Am/Ph sheath cells
Intestine
Somatic gonad precursors
Seam cells
Non-seam hypodermis
Distal tip cells
Pharyngeal epithelia
Coelomocytes
Body wall muscle
Germline
Vulval precursors
Sex myoblasts
Excretory cells
Pharyngeal muscle
Rectum
Intestinal/rectal muscle
Pharyngeal gland
TF binding to the ciliome
G
FKH-8
ciliome genes bound by the TF
446 TF ChiPseq data sets
H
FKH-8
ciliome genes bound by the TF
Total number of peaks

Figure 1

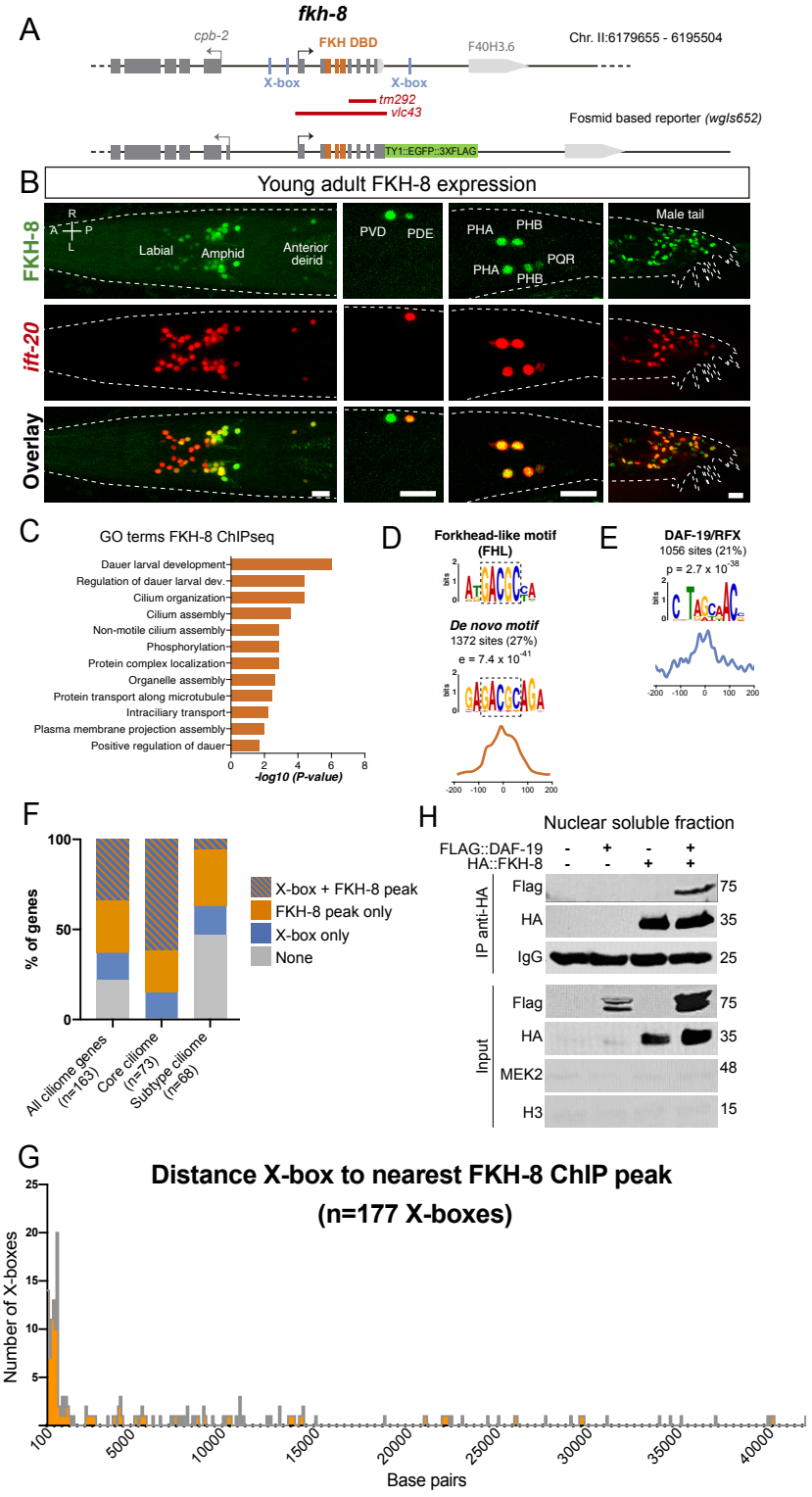
A
fkh-8
cpb-2
FKH DBD
F40H3.6
Chr. II:6179655 - 6195504
X-box
X-box
tm292
vlc43
Fosmid based reporter (wgIs652)
TY1::EGFP::3XFLAG
B
Young adult FKH-8 expression
FKH-8
ift-20
Overlay
Labial
Amphid
Anterior deirid
PVD
PDE
PHA
PHB
PQR
Male tail
C
GO terms FKH-8 ChIPseq
Dauer larval development
Regulation of dauer larval dev.
Cilium organization
Cilium assembly
Non-motile cilium assembly
Phosphorylation
Protein complex localization
Organelle assembly
Protein transport along microtubule
Intraciliary transport
Plasma membrane projection assembly
Positive regulation of dauer
-log10 (P-value)
D
Forkhead-like motif (FHL)
De novo motif
1372 sites (27%)
E
DAF-19/RFX
1056 sites (21%)
F
% of genes
X-box + FKH-8 peak
FKH-8 peak only
X-box only
None
All ciliome genes (n=163)
Core ciliome (n=73)
Subtype ciliome (n=68)
H
Nuclear soluble fraction
FLAG::DAF-19
HA::FKH-8
IP anti-HA
Input
Flag
HA
IgG
MEK2
H3
G
Distance X-box to nearest FKH-8 ChIP peak
(n=177 X-boxes)
Number of X-boxes
Base pairs

Figure 2

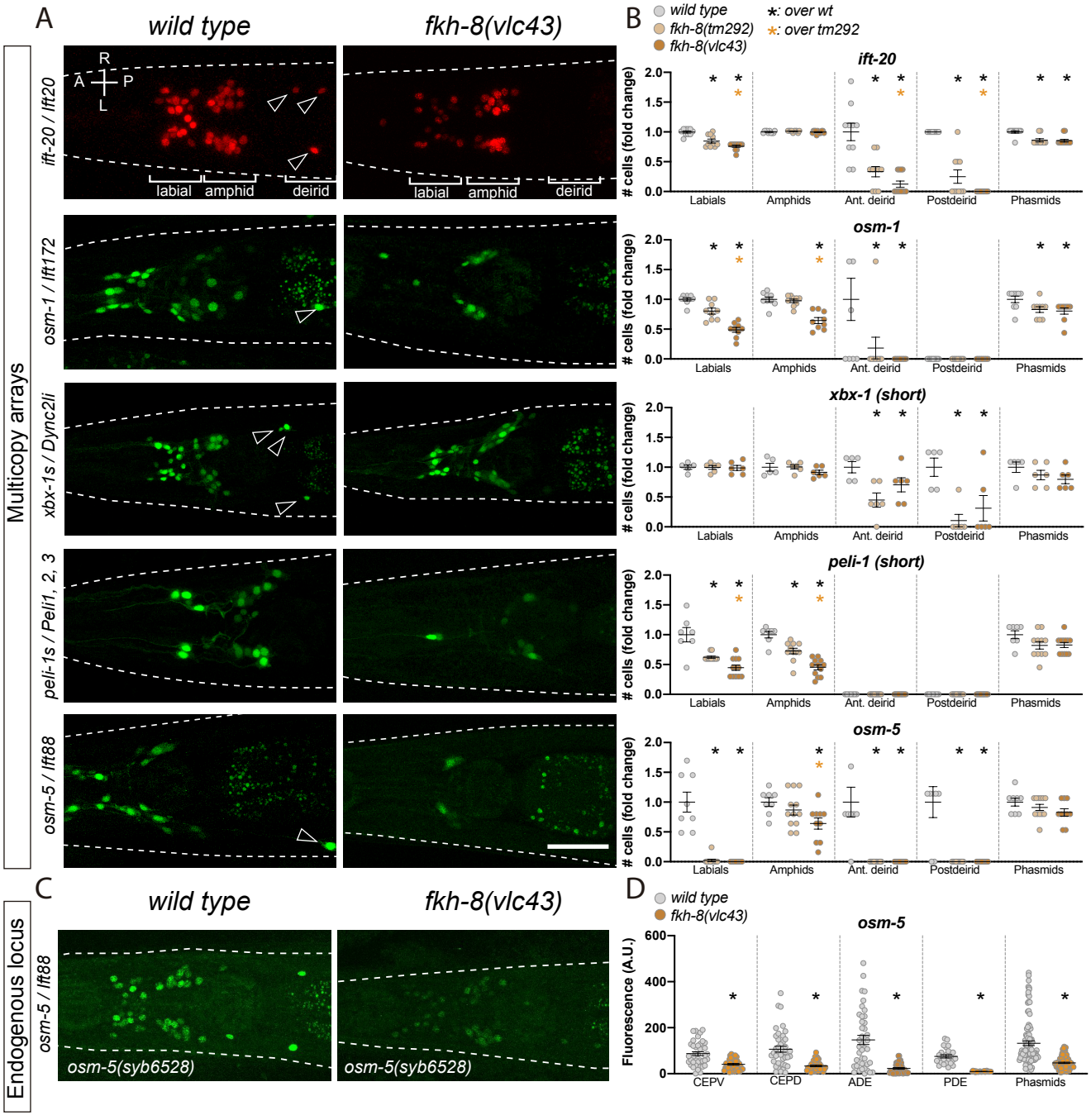
A
wild type
fkh-8(vlc43)
Multicopy arrays
ift-20 / Ift20
osm-1 / Ift172
xbx-1s / Dync2li
peli-1s / Peli1, 2, 3
osm-5 / Ift88
R
A
P
L
labial
amphid
deirid
B
wild type
fkh-8(tm292)
fkh-8(vlc43)
*: over wt
*: over tm292
ift-20
osm-1
xbx-1 (short)
peli-1 (short)
osm-5
cells (fold change)
Labials
Amphids
Ant. deirid
Postdeirid
Phasmids
C
Endogenous locus
osm-5 / Ift88
osm-5(syb6528)
D
Fluorescence (A.U.)
CEPV
CEPD
ADE
PDE

Figure 3

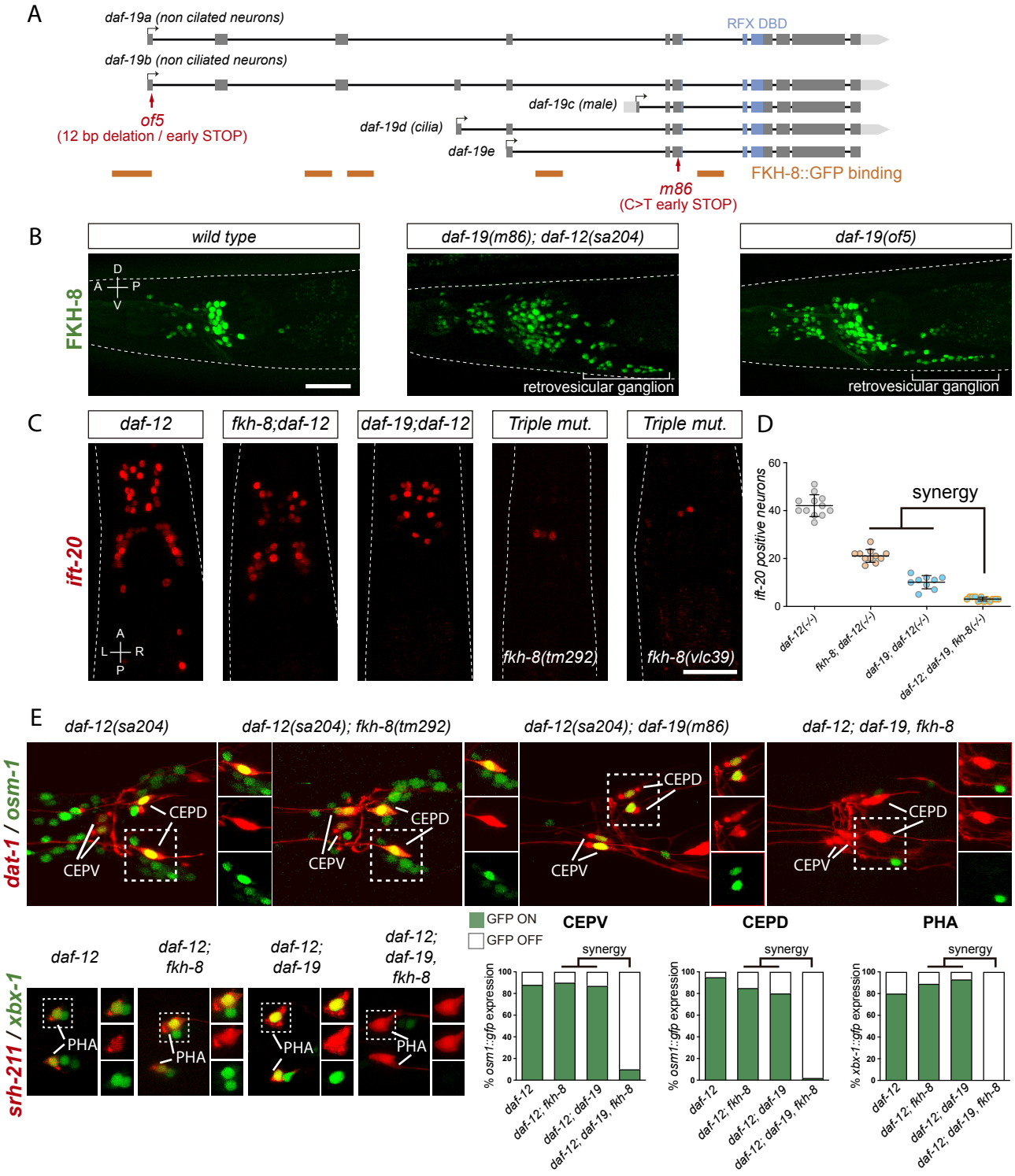
A
daf-19a (non cilated neurons)
RFX DBD
daf-19b (non ciliated neurons)
of5
(12 bp delation / early STOP)
daf-19c (male)
daf-19d (cilia)
daf-19e
FKH-8::GFP binding
m86
(C>T early STOP)
B
wild type
daf-19(m86); daf-12(sa204)
daf-19(of5)
FKH-8
retrovesicular ganglion
C
daf-12
fkh-8;daf-12
daf-19;daf-12
Triple mut.
Triple mut.
ift-20
fkh-8(tm292)
fkh-8(vlc39)
D
ift-20 positive neurons
synergy
daf-12(-/-)
fkh-8; daf-12(-/-)
daf-19; daf-12(-/-)
daf-12; daf-19, fkh-8(-/-)
E
daf-12(sa204)
daf-12(sa204); fkh-8(tm292)
daf-12(sa204); daf-19(m86)
daf-12; daf-19, fkh-8
dat-1 / osm-1
CEPD
CEPV
srh-211 / xbx-1
daf-12
daf-12; fkh-8
daf-12; daf-19
daf-12; daf-19, fkh-8
PHA
GFP ON
GFP OFF
CEPV
CEPD
PHA
synergy
% osm1::gfp expression
% xbx-1::gfp expression

Figure 4

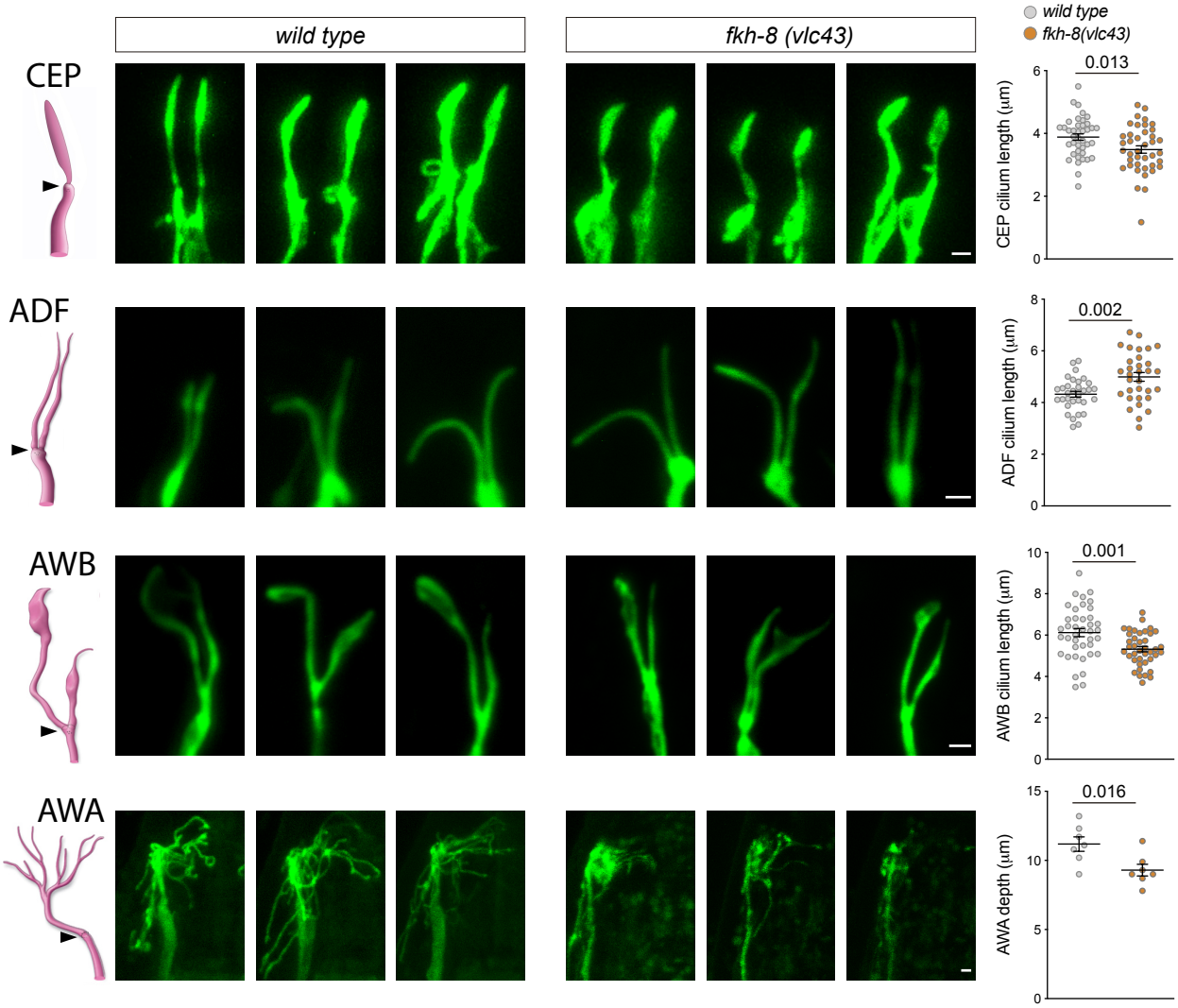
wild type
fkh-8 (vlc43)
CEP
ADF
AWB
AWA
wild type
fkh-8(vlc43)
0.013
CEP cilium length (µm)
0.002
ADF cilium length (µm)
0.001
AWB cilium length (µm)
0.016
AWA depth (µm)

Figure 5

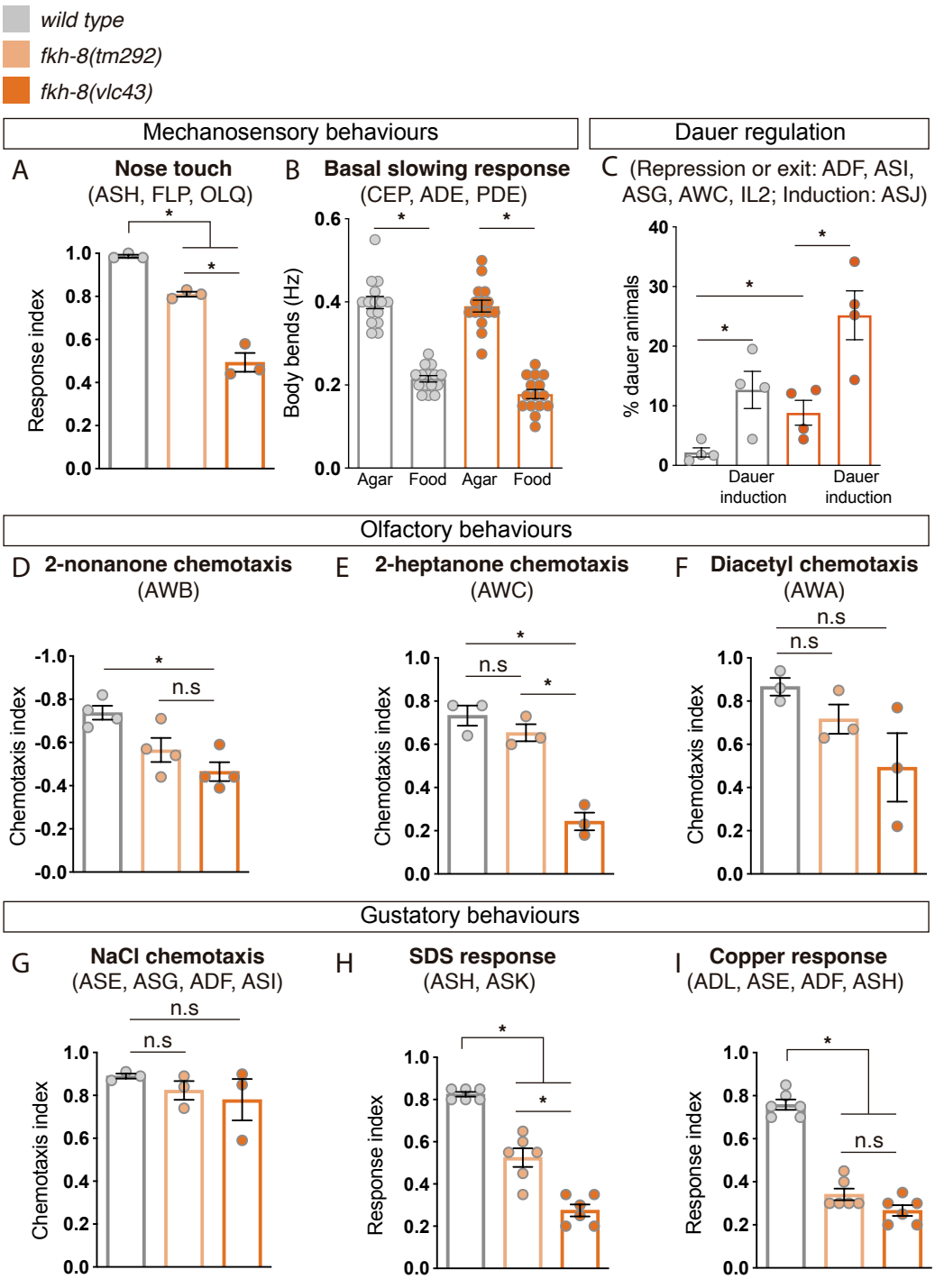
wild type
fkh-8(tm292)
fkh-8(vlc43)
Mechanosensory behaviours
Dauer regulation
A
Nose touch
(ASH, FLP, OLQ)
Response index
B
Basal slowing response
(CEP, ADE, PDE)
Body bends (Hz)
Agar
Food
Agar
Food
C
(Repression or exit: ADF, ASI, ASG, AWC, IL2; Induction: ASJ)
% dauer animals
Dauer induction
Dauer induction
Olfactory behaviours
D
2-nonanone chemotaxis
(AWB)
Chemotaxis index
n.s
E
2-heptanone chemotaxis
(AWC)
Chemotaxis index
n.s
F
Diacetyl chemotaxis
(AWA)
Chemotaxis index
n.s
n.s
Gustatory behaviours
G
NaCl chemotaxis
(ASE, ASG, ADF, ASI)
Chemotaxis index
n.s
n.s
H
SDS response
(ASH, ASK)
Response index
I
Copper response
(ADL, ASE, ADF, ASH)
Response index
n.s

Figure 6

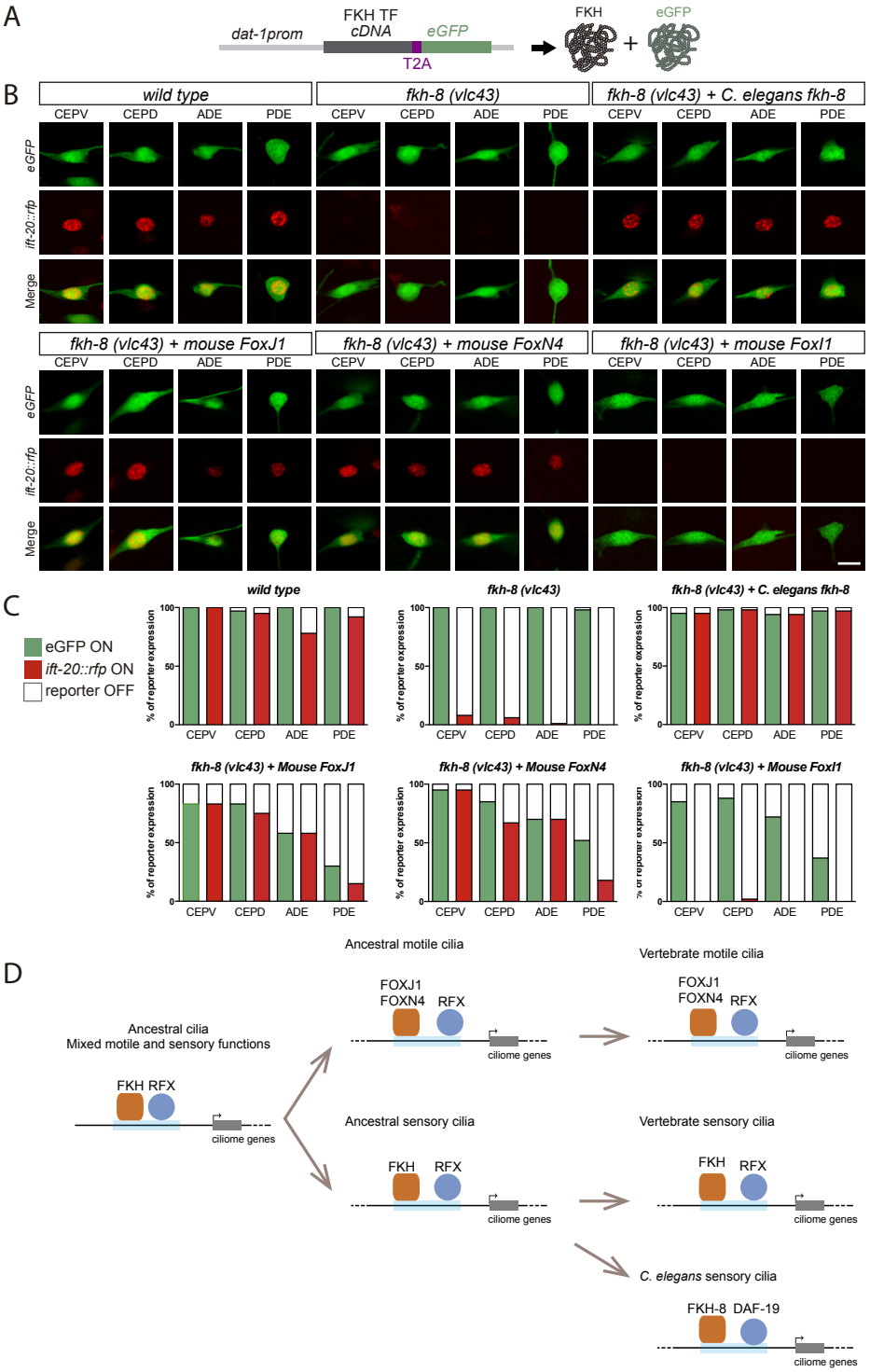
A
dat-1prom
FKH TF
cDNA
eGFP
T2A
FKH
eGFP
B
wild type
fkh-8 (vlc43)
fkh-8 (vlc43) + C. elegans fkh-8
fkh-8 (vlc43) + mouse FoxJ1
fkh-8 (vlc43) + mouse FoxN4
fkh-8 (vlc43) + mouse Foxl1
CEPV
CEPD
ADE
PDE
eGFP
ift-20::rfp
Merge
C
eGFP ON
ift-20::rfp ON
reporter OFF
% of reporter expression
D
Ancestral cilia
Mixed motile and sensory functions
FKH RFX
ciliome genes
Ancestral motile cilia
FOXJ1
FOXN4 RFX
Vertebrate motile cilia
Ancestral sensory cilia
Vertebrate sensory cilia
C. elegans sensory cilia
FKH-8 DAF-19

Figure 7

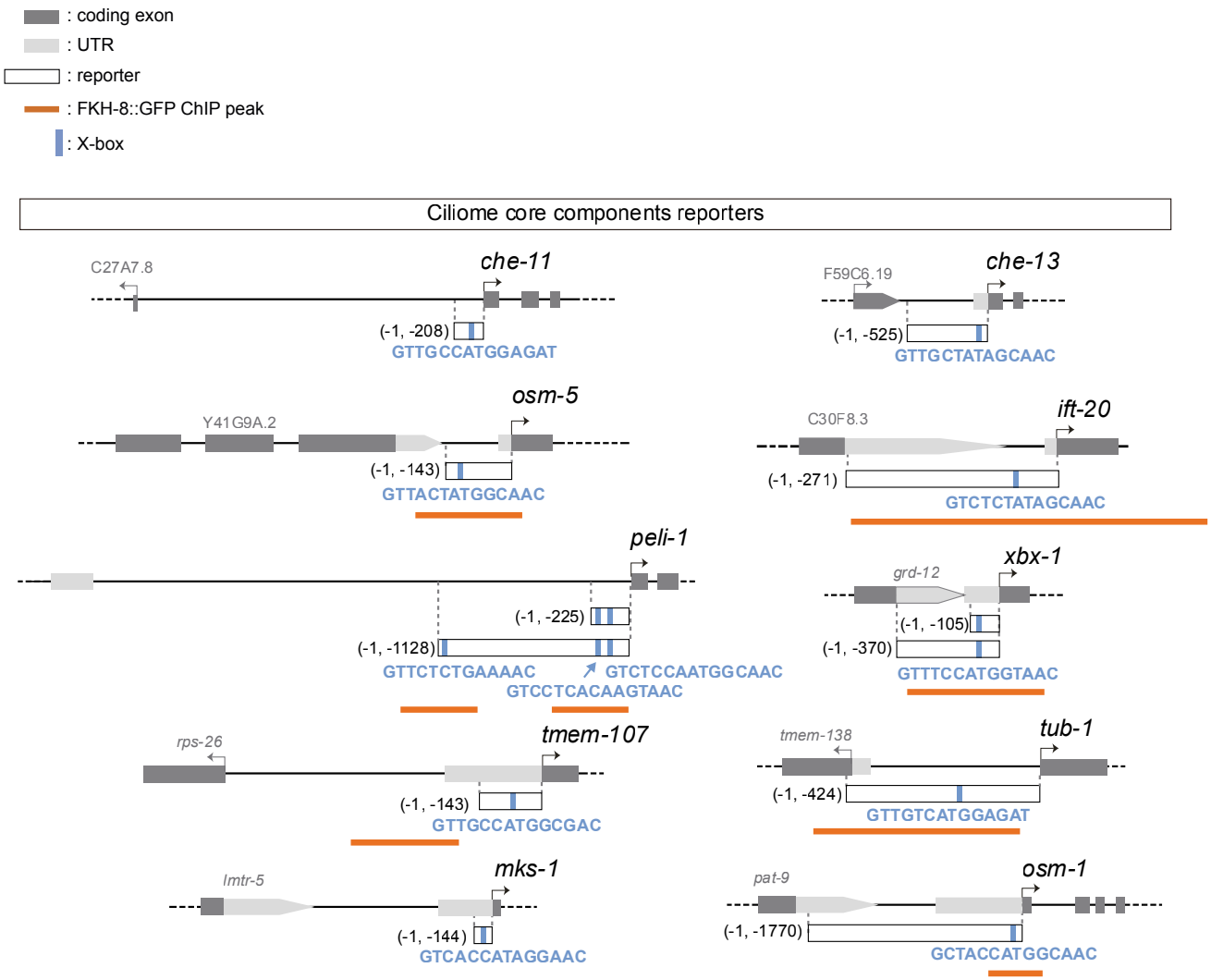

Figure 1 Figure Supplement 1

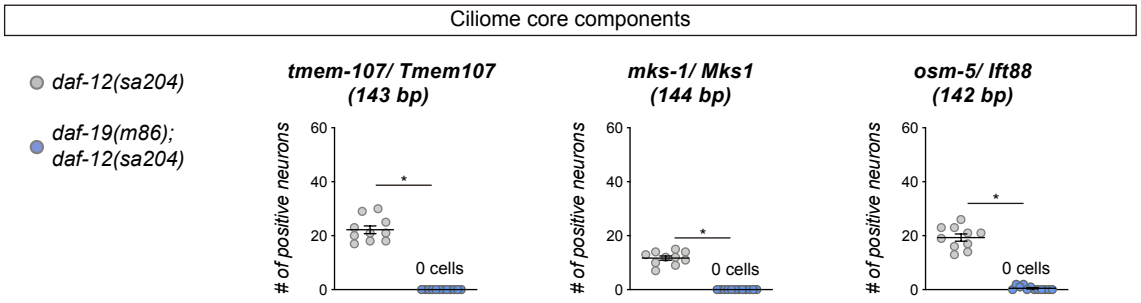

Figure 1 Figure Supplement 2

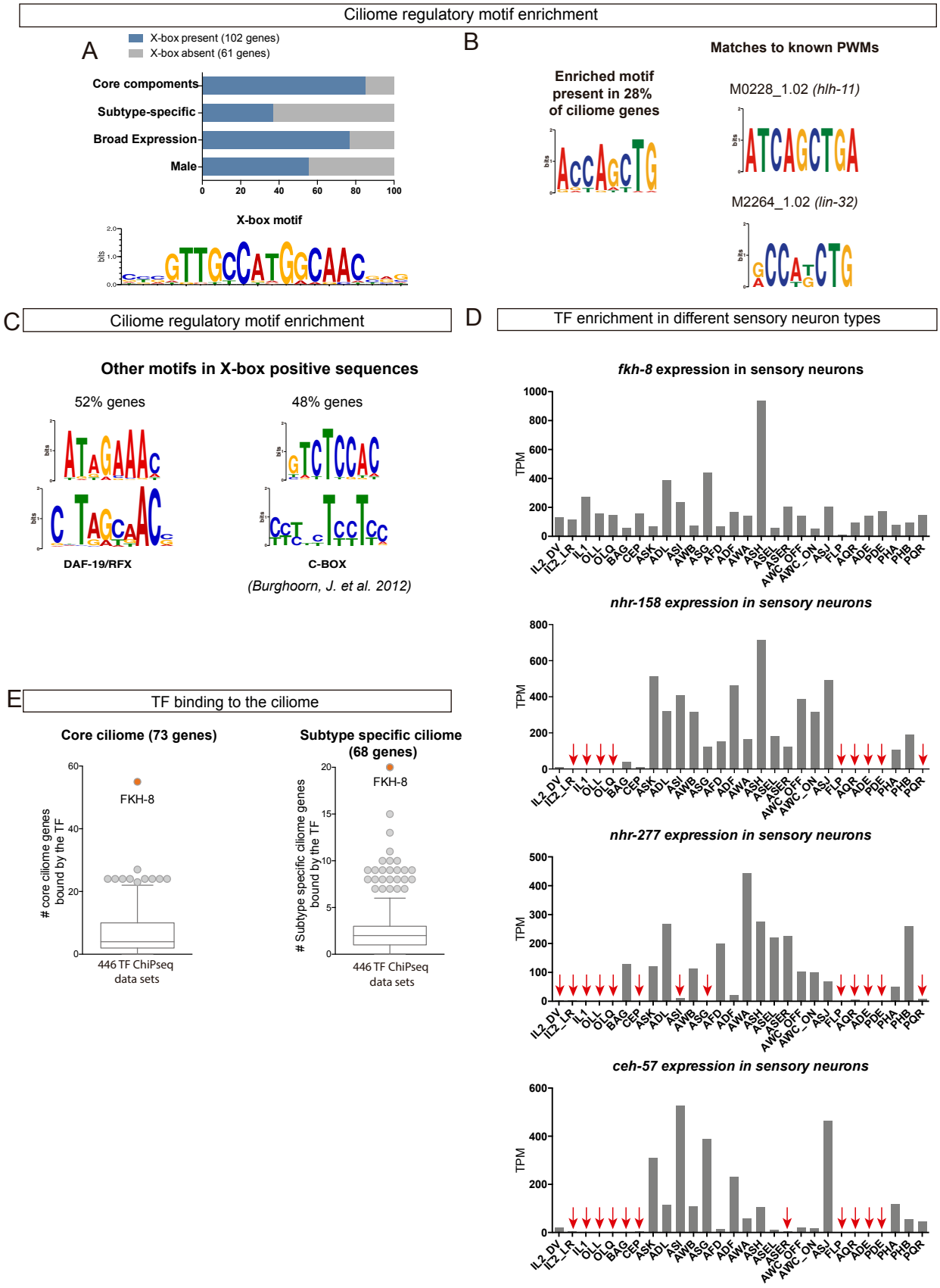

Figure 1 Figure Supplement 3

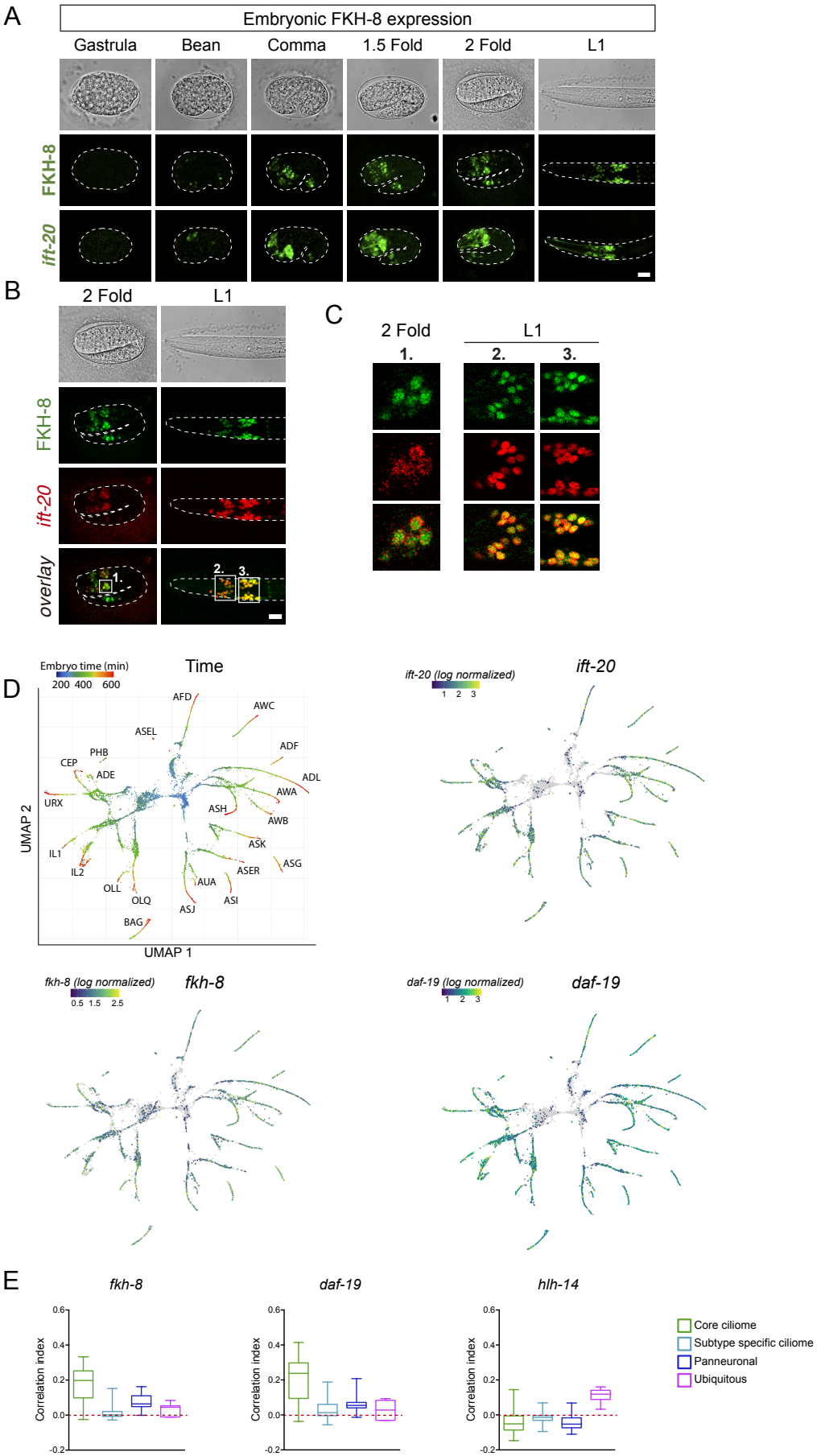

Figure 2 Figure Supplement 1

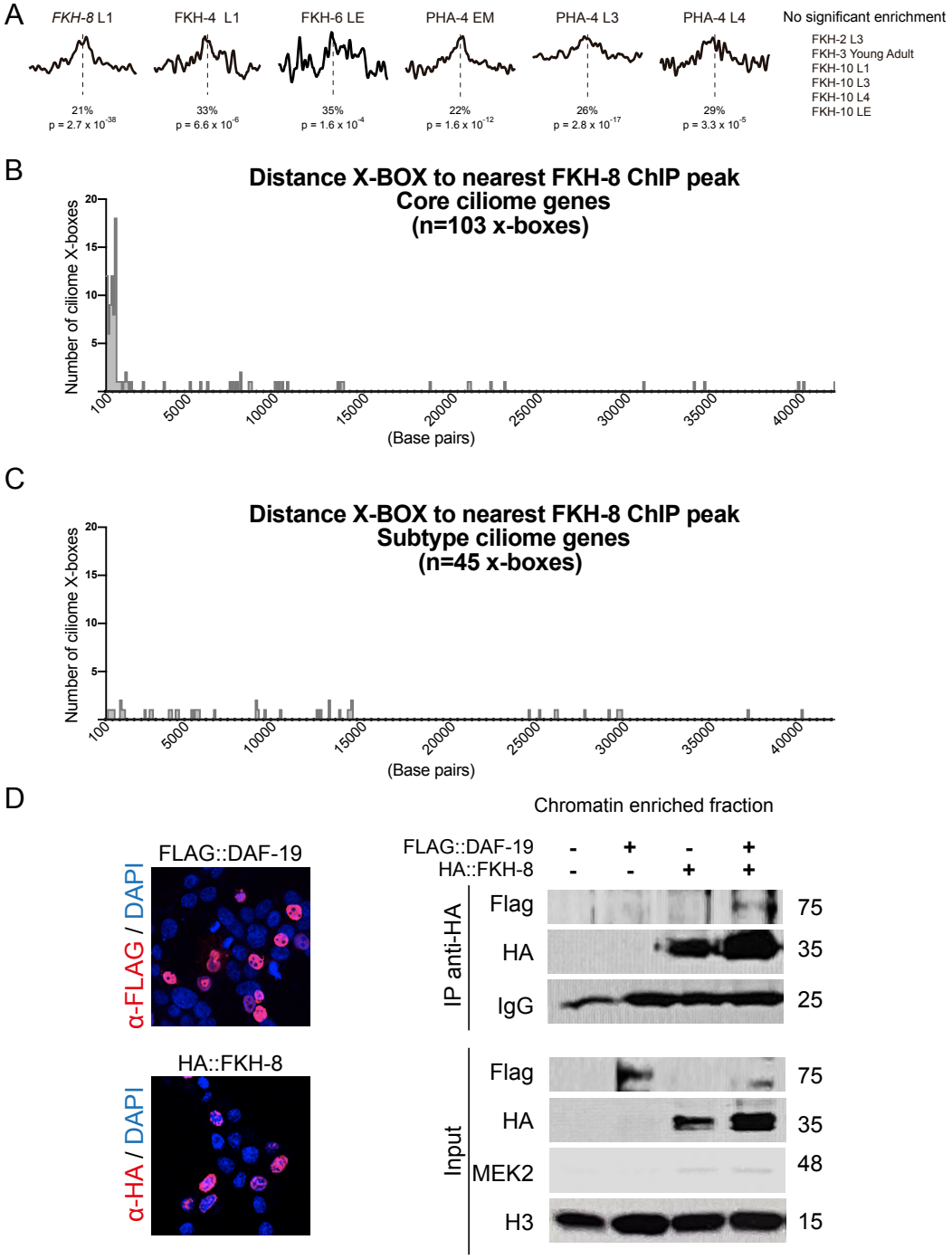

Figure 2 Figure Supplement 2

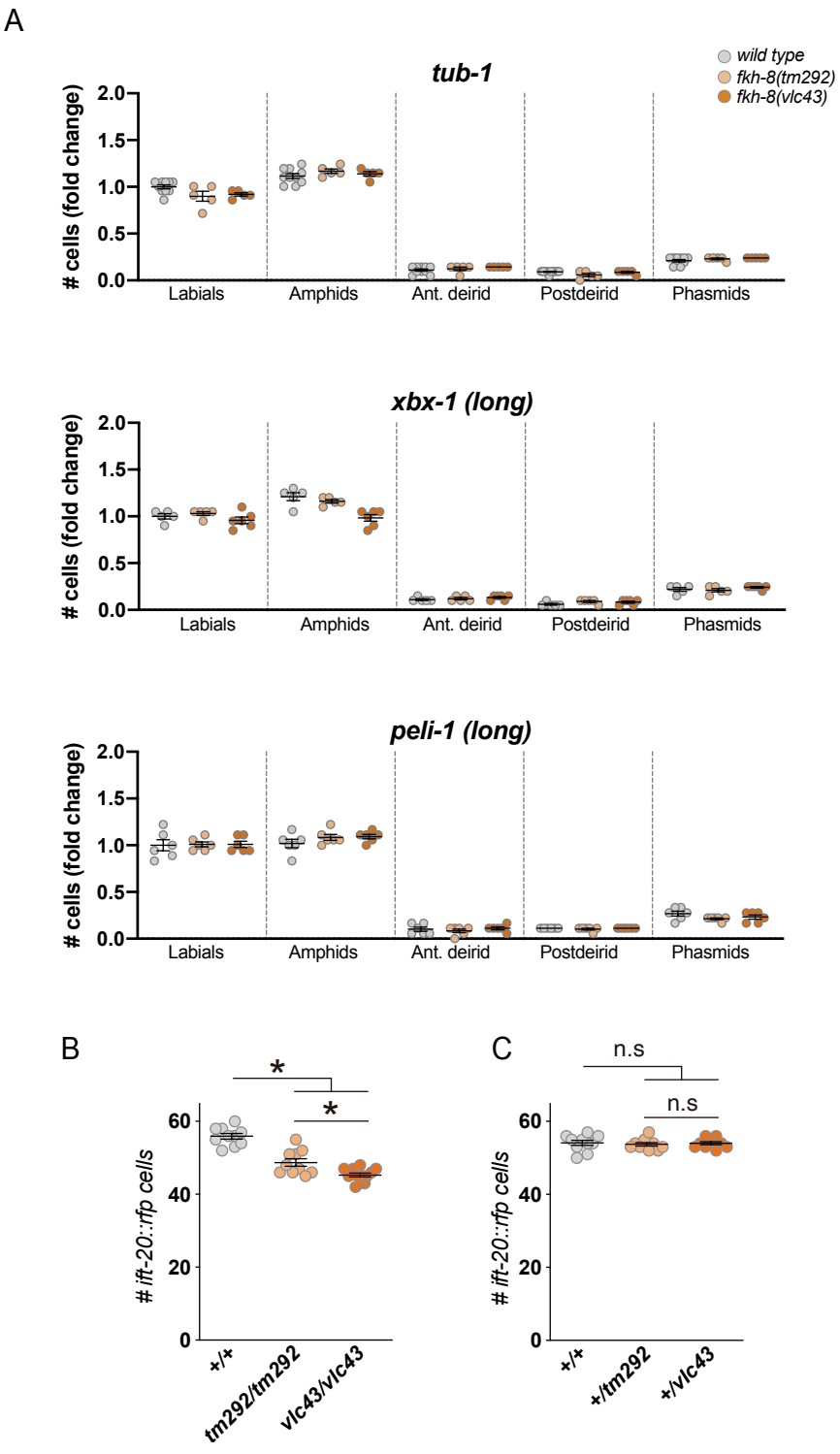

Figure 3 Figure Supplement 1

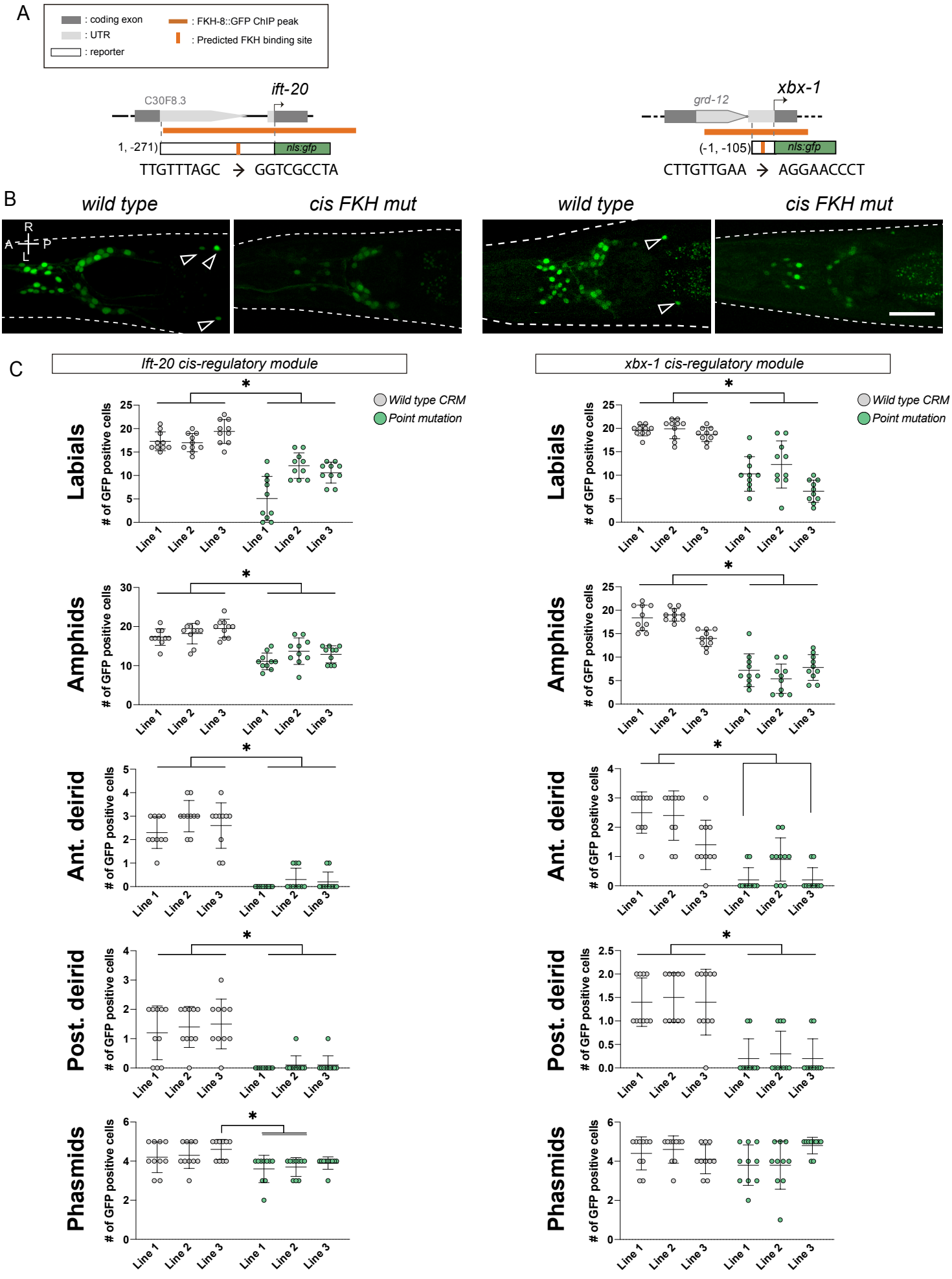

Figure 3 Figure Supplement 2

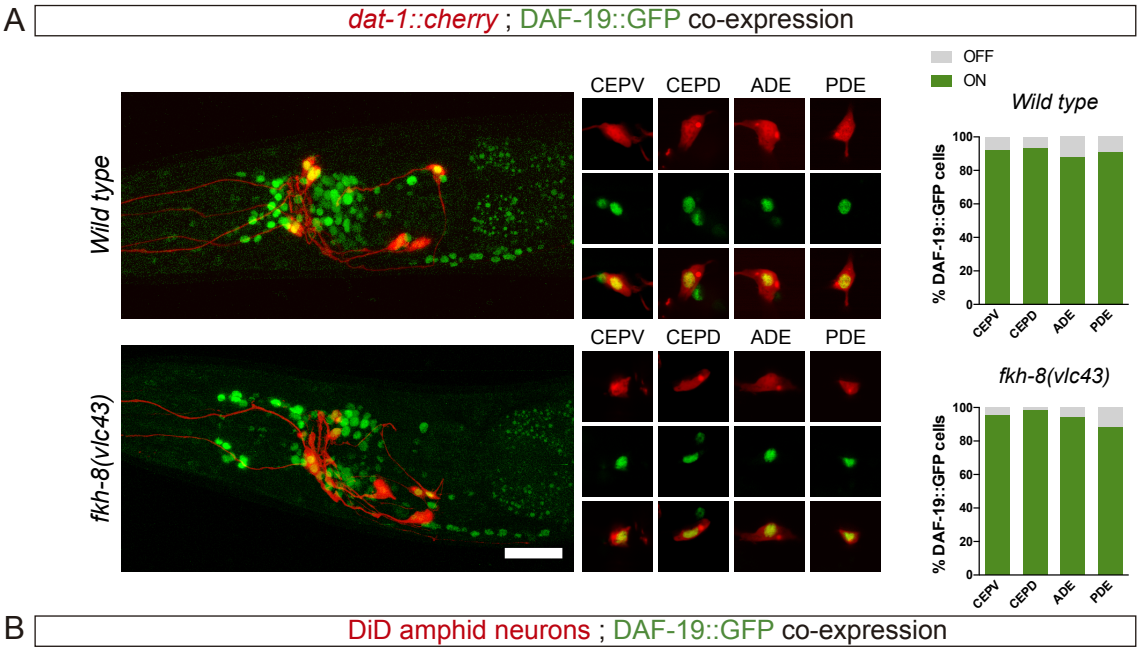

Figure 4 Figure Supplement 1

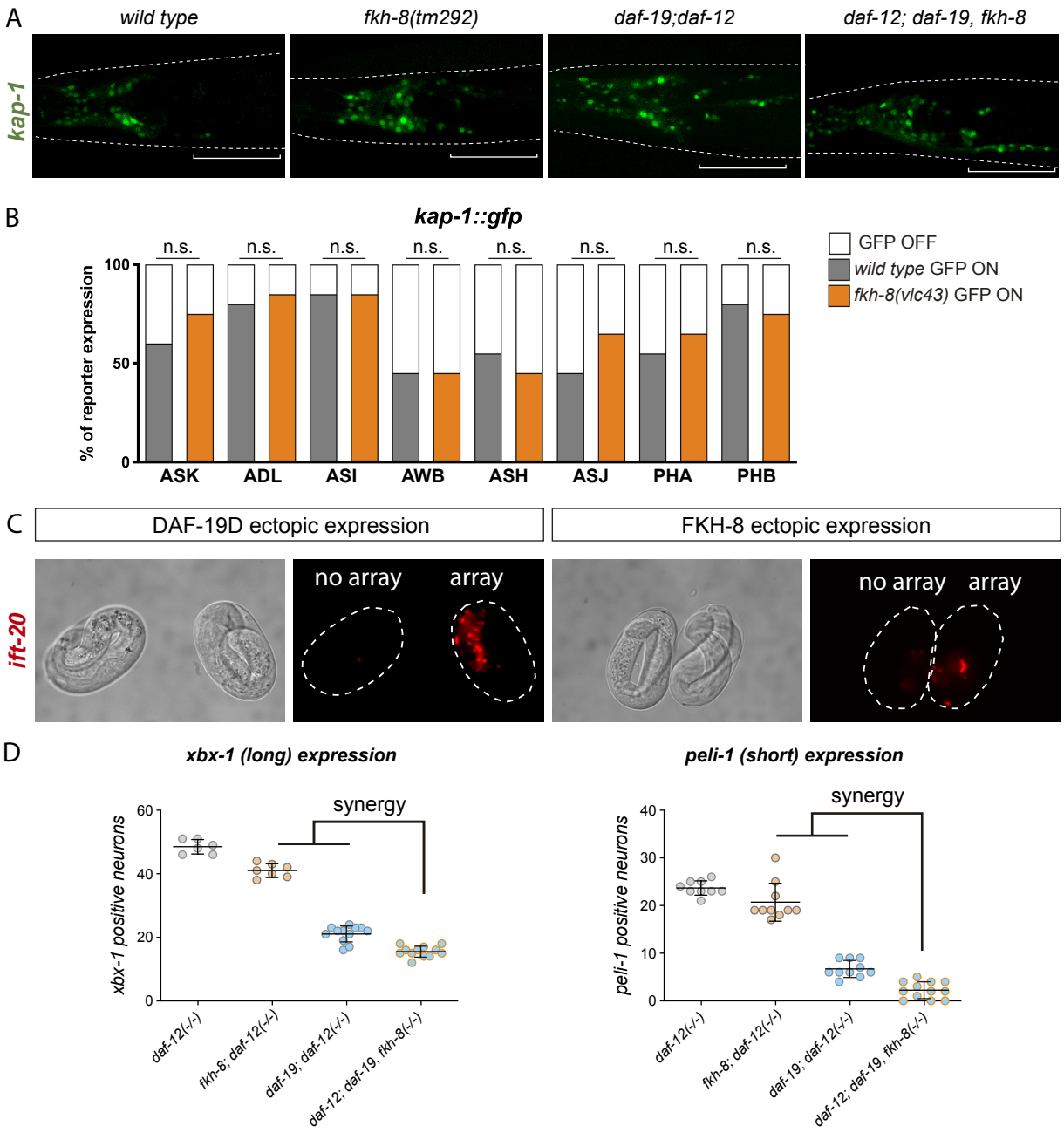

Figure 4 Figure Supplement 2

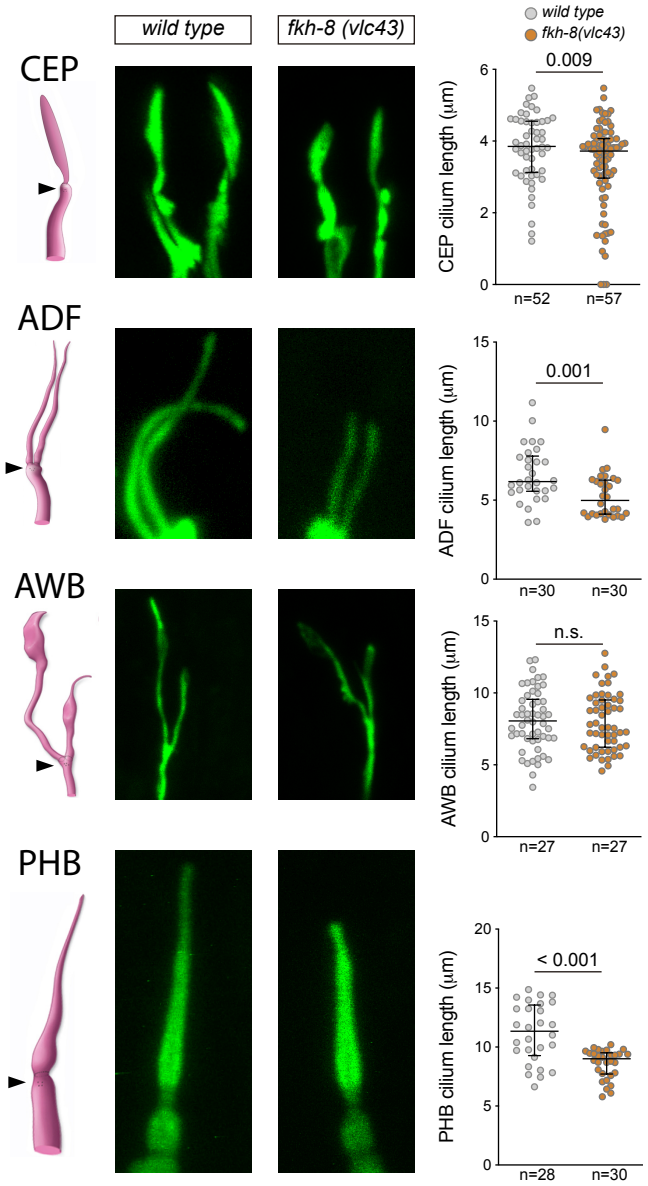

Figure 5 Figure Supplement 1

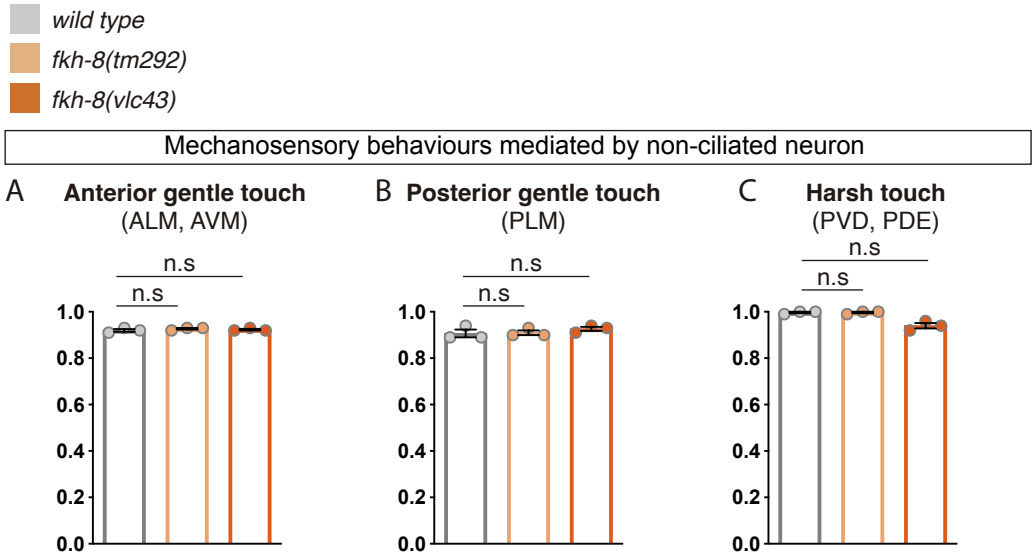

Figure 6 Figure Supplement 1