Transcriptional regulation of the ciliopathy gene *MKS1*/*mks-1*

by

Ting Zhang

B.Sc., Fudan University, 2011

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name: **Ting Zhang**

Degree: **Master of Science**

Title of Thesis: **Transcriptional Regulation of the ciliopathy gene *MKS1/mks-1***

Examining Committee: **Chair:** Ryan Morin
Assistant Professor

Jack Chen
Senior Supervisor
Professor

David Baillie
Supervisor
Professor

Nancy Hawkins
Supervisor
Associate Professor

Harald Hutter
Internal Examiner
Professor
Department of Biological Sciences

Date Defended: June 12, 2014

Partial Copyright Licence

Abstract

Meckel Grubber Syndrome (MKS) is a severe ciliopathy. The first identified causal loci of MKS resided in the gene *MKS1*. The only known transcription factor of *MKS1* is Regulatory Factor X (RFX). *C. elegans*, an organism with well-characterized ciliated neurons, is ideal for studying ciliary genes. Many ciliary genes including *MKS1* are conserved in *C. elegans*. My study aims to find candidate mutants of transcription factor(s) for *mks-1*, the *C. elegans* ortholog of *MKS1*. In order to track in vivo *mks-1* expression in *C. elegans*, I generated a transgenic strain that expressed Green Fluorescence Protein (GFP) driven by *mks-1* promoter. I carried out genetic screens to identify mutations that altered GFP fluorescence intensity and expression profile, as a way to search for potential transcription factor(s) for *mks-1* in *C. elegans*. I successfully found an X-linked recessive mutant, which suppresses *mks-1* reporter gene expression in subset of labial neurons. The mutant has normal labial neurons development. This research has set up a solid stage for studying the transcriptional regulation of *mks-1* and other ciliary genes.

Keywords: Transcription factor; *MKS1/mks-1*; Mutagenesis; Genetic Screening

Dedication

To my family, my parents and my husband. They enlightened my way in the darkness.

To my best friends. I was lucky to have met them on the campus of Fudan University.

To my dreams. My dreams led me to the path of honesty and courage.

To my heart. I followed my heart faithfully all the time.

To my future. I am waiting for a promising and bright future.

Acknowledgements

Most importantly, I would like to thank my senior supervisor Dr. Jack Chen, who kindly gave me the opportunity to study in Simon Fraser University, who generously supported my thesis work, and who nicely advised me through my three-year master's work. I would like to thank my supervisory committee members Dr. David Baillie and Dr. Nancy Hawkins. I would like to thank them for their important advice and support on my project. I would like to thank all colleagues in Chen lab in general: Jun Wang for training me when I first came to Chen lab, for preparing the GFP PCR fusion construct of transcriptional *mks-1* expression, and for designing some primers in my thesis work; Zhaozhao Qin for kindly injecting the transcriptional *mks-1* expression constructs into *C. elegans*, for help on my first round of genetic screening, and for helping with X-ray integrations of transgenic strains; Dr. Maja Tarailo-Graovac for her advice on my experiments; Dr. Jeffery Chu for his help on genomic DNA preparation and whole-genome sequencing; Timothy Warrington for suggestions on my work and pointing out some grammar errors in my thesis draft; Judy Kim, a summer NSERC-USRA undergraduate student, for help on second round of genetic screening; Dr. Christian Frech, Tammy Wong, Dr. Jiarui Li for their good suggestions in my work. It was interesting and inspiring to work with Chen lab colleagues. I would also like to thank members from Baillie lab, especially Shuyi Chua, for her helpful suggestions in my troubleshooting. I would like to thank Dr. Robert Johnson from Baillie lab for providing some strains used in this study. I would like to thank Tim Heslip for training me on confocal microscopy performance and helping me with confocal microscopy issues. Some strains were hosted by the CGC, which is funded by NIH Office of Research infrastructure Programs (P40OD010440). Also some strains were obtained from NBRP. Last but not least, I would like to thank all the friends I met in SFU. It was encouraging to discuss science with friends in MBB. The three-year research experience in MBB would be a turning point in my life, so I am appreciated that I have stayed here and learned a lot.

Table of Contents

List of Tables

List of Figures

List of Acronyms

BBS	Bardet-biedl Syndrome
CRX	Cone-rod Homeobox protein
EMS	Ethyl methanesulfonate
FOX	Forkhead Box protein
GFP	Green Fluorescence Protein
IFT	Intraflagellar transport
JBTS	Joubert Syndrome
MKS	Meckel Syndrome
NPHP	Nephronophthisis
RFX	Regulatory Factor X
tdTomato	Tandem dimer Tomato protein
VHL	Von Hippel–Lindau Disease

Glossary

Bardet-biedl Syndrome	One type of ciliopathy. It is a mild ciliopathy. Major symptoms are obesity, retinitis pigmentosa, polydactyly, and hypogonadism.
Cone-red Homeobox protein	A photoreceptor-specific transcription factor
Ethyl methanesulfonate	An alkylating mutagen that causes GC to AT transitions
Forkhead Box protein	Transcription factor that regulates gene expressions in cell differentiation, growth, proliferation and longevity
Green Fluorescence Protein	A protein that shows green fluorescence in blue ultraviolet light. Isolated from jellyfish.
Intraflagellar transport	Bidirectional mobility along axonemal microtubules in cilia formation
Joubert Syndrome	One type of ciliopathy. It is a moderate ciliopathy. Major symptoms are ataxia and hyperpnea.
Meckel Syndrome	One type of ciliopathy. It is a severe and lethal ciliopathy. Major symptoms are renal cystic dysplasia, encephalocele, polydactyly, pulmonary hypoplasia and hepatic development defects.
Nephronophthisis	One type of ciliopathy. It is a mild ciliopathy. It causes problems in patient's urinary system.
Regulatory Factor X	Transcription factor that regulates ciliary genes
Tandem dimer Tomato protein	Proteins that show exceptionally red fluorescence under a 581 nm wavelength. Developed by Dr. Roger Tsien
Von Hippel–Lindau Disease	One type of ciliopathy. It is a moderate ciliopathy. Patients have tumors in central nervous system or retina.

Chapter 1.

General Introduction

1.1. *C. elegans*

Caenorhabditis elegans is a nematode species that Sydney Brenner proposed to use as a model organism for studying neuronal development in the 1960s. He carried out an unprecedented genetic analysis of *C. elegans,* which set up a platform for many ground breaking scientific discoveries (Brenner 1974). An adult hermaphrodite *C. elegan*s has 959 somatic cells, 302 of which are neurons (Sulston and Horvitz 1977). An adult male *C. elegans* has 1031 somatic cells, 381 of which are neurons (Sulston *et al*. 1980). The simple body plan and well-defined cell lineage make *C. elegans* an ideal organism for genetics studies (Brenner 1974). *C. elegans* has two genders, males and hermaphrodites, based on an X-O sex determination system. The males contain one X chromosome, while the hermaphrodites have two copies. Researchers are able to maintain *C. elegans* easily in a lab environment, because of its simple sex determination system and self-reproducing ability of hermaphrodites. An adult *C. elegans* is around 1 mm in length (Figure 1.1). In Brenner's studies, he screened for *C. elegans* mutant phenotypes, including uncoordinated, rolling, long, dumpy, blister, and lethal (Brenner 1974). Since then, *C. elegans* has been widely used as a model organism for neuroscience and genomics studies. *C. elegans* has a small genome size. The 98 megabytes genome sequence of *C. elegans* is first published in 1998, with some gaps left at that time. More than 19,000 protein-coding gene sequences were identified (Consortium 1998). The publication of *C. elegans* genome information accelerates genomics studies on *C. elegans*. Now with persistent efforts, *C. elegans* is the only animal whose genome has been completely deciphered. The whole *C. elegans* genome is 100,291,840bp including mitochondrial genome (Hillier *et al.* 2005).

At 20°C, it takes 3.5 days for wildtype *C. elegans* to develop from embryos to adults. The worms go through four larvae stages, L1-L4, during development (Figure 1.2).

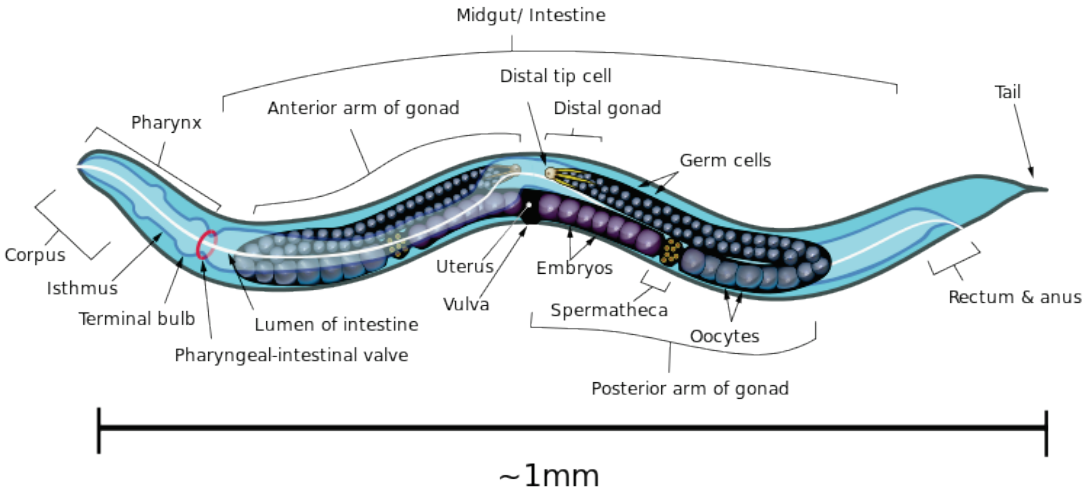


Figure 1.1. Anatomical structure of *C. elegans*
Note. Figure source:
http://en.wikipedia.org/wiki/File:Caenorhabditis_elegans_hermaphrodite_adult-en.svg

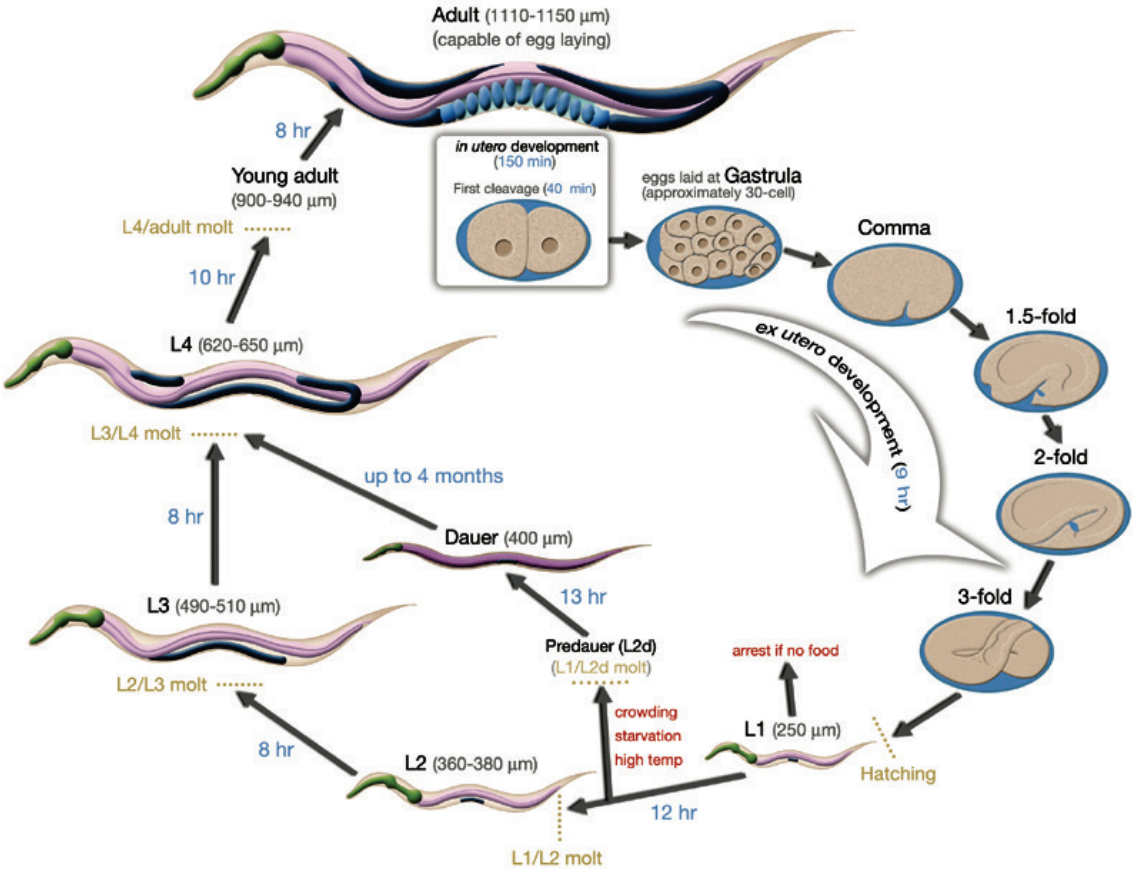


Figure 1.2. Life cycle of *C. elegans*
Note. Life cycle of *C. elegans* at 22°C. Figure source:
http://www.wormatlas.org/ver1/handbook/anatomyintro/anatomyintro.htm

1.2. Cilia and Ciliogenesis

The word “cilia” means hair or eyelash in Latin. The cilia were first observed by Antoni van Leeuwenhoek more than 300 years ago (Leeuwenhoek 1677). He described cilia as thin legs for protozoa that move them nimbly (Leeuwenhoek 1677; Bloodgood 2009). The term cilium was first used by Muller in 1786 (Muller 1786). Cilia were defined as “active organelles moved by contractile materials along their length” in a review in 1835 (Sharpey 1835; Gibbons 1981), but now cilia are defined as small hair-like organelles that are projected from eukaryotic cell surfaces (Goetz and Anderson 2010).

Now scientists know there are two types of cilia: motile and non-motile cilia (Satir and Christensen 2008). In 1950s, scientists were able to observe cilia structures aided

with Transmission Electron Microscopy techniques. A cilium is divided into three parts: basal body, middle segment and distal segment. The components of a cilium are a ring of microtubule scaffolding called the axoneme and the plasma membrane. Primary cilia have “9+0” microtubule structures (Sotelo and Trujillo-Cenoz 1958). It is called a “9+0” structure because the axoneme in primary cilia lacks the central microtubules. Motile cilia have a “9+2” microtubule structure, because the axoneme has a pair of central microtubules (Sotelo and Trujillo-Cenoz 1958). In cilia structure, “9” stands for 9 microtubules. The axoneme consists of 9 doublet microtubules surrounding the periphery in the middle segment, and 9 singlet microtubules surrounding the periphery in the distal segment (Fisch and Dupuis-Williams 2011). The axoneme is extended from the basal body, which develops from the centriole after cytokinesis, and which connects the cilia to the cell body. In mammals’ ciliated cells, the axoneme consists of 9 triplet microtubules surrounding the periphery in the basal body, but the triplet microtubules degenerate in the basal body of *C. elegans’* ciliated cells (Perkins *et al*. 1986).

Motile cilia help cells move themself, or move fluids and dirt away from the surface of cells. Non-motile cilia, also named primary cilia, are necessary for cells to sense the environment and to work in signal transduction. Cells with primary cilia are found in many sensory organs in humans. For example, primary cilia are found on the surface of photoreceptor cells in the retina and olfactory cells in the nose. Primary cilia are also very important for kidney cells. Praetorius and Spring observed the mechanoreceptive ion channels (TRP channels) in primary cilia on mammalian kidney cell membranes(Praetorius and Spring 2003). Primary cilia are important in many signalling pathways for vertebrate development. Studies have demonstrated that primary cilia play roles in Hedgehog signalling (i.e. Sonic Hedgehog signalling in mammals) and Notch signalling (Goetz and Anderson 2010). Ezratty *et al.* discovered the role of primary cilia in epidermal differentiation of skin development: they are primarily involved in Notch signalling (Ezratty *et al.* 2011).

Studies in cilia have been boosted since 1990s. Because cilia are very conserved organelles in eukaryotes, many model organisms have been used in the studies of primary cilia, including *Chlamydomonas, C. elegans*, *Drosophila* and mammalian cells.

C. elegans is an ideal model organism to study the function of primary cilia in sensory neurons. Of 302 sensory neurons in *C. elegans*, 60 are ciliated sensory neurons. These ciliated neurons include amphids, phasmids, inner labial, outer labial and some other types (Figure 1.3). The amphids, labial neurons and phasmids appear in bilaterally symmetric pairs (Bargmann 2006). There are 12 pairs of amphids ciliated neurons, named ASH, ASI, ASG, ASE, ASJ, ASK, ADF, ADI, AWA, AWB, AWC and AFD (Ward *et al*. 1975; Perkins *et al*. 1986), and two pairs of phasmids ciliated neurons, named PHA and PHB (Hall and Russell 1991). There are two kinds of inner labial neurons: IL1, IL2, with each consisting of three pairs of neurons (Ward *et al*. 1975; Ware 1975). In the outer labial, there are two pairs of OLL neurons and four pairs of OLQ neurons (Ward *et al*. 1975; Ware 1975).

C. elegans primary cilia possess the "9+0" structure. The primary cilium has three components: the base, which is a degenerated basal body that consists of a structure termed transition zone, the middle segment, which has 9 doublets microtubules, and the distal segment, which has 9 singlet microtubules (Figure 1.4) (Gilula and Satir 1972; Evans *et al*. 2006). *C. elegans* ciliary base structure is different from other species. The axoneme is degenerated in basal body, but still the basal body contains transition fibres (Perkins *et al*. 1986). Axoneme structure of primary cilia in *C. elegans* is same as that in other species. The conserved cilia structures in *C. elegans* make it ideal for cilia studies.

All cilia detected in *C. elegans* are primary cilia. Defective cilia cause a variety of phenotypes in *C. elegans*, an indication of functional failure of cilia. The major phenotype is failure to take DiI or DiO in the 6 amphids and 2 phasmids neurons (Dyf phenotypes) (Herman 1984; Hedgecock *et al*. 1985). Mutants with defective cilia also have these phenotypes: not able to avoid regions with high osmotic strength (Osm phenotype) (Culotti and Russell 1978), unable to perform chemotaxis (Che phenotype) (Bargmann *et al*. 1993), and reduced ability to respond to touch (Mec phenotype) (Chalfie and Sulston 1981). Some ciliary mutants might have abnormal male mating behaviour (Barr and Sternberg 1999), constitutive dauer formation (Riddle *et al*. 1981), or abnormal response to noxious temperatures (Wittenburg and Baumeister 1999). The ciliary gene *tub-1* mutant is connected to abnormal lipid accumulation (Mak *et al*. 2006).

Ciliogenesis starts after cytokinesis. The centriole migrates, and then fuses to the plasma membrane. The ciliary vesicles are docked to the mother centriole to form the basal body (Sorokin 1962; Sorokin 1968). Meanwhile, the cilia membrane is formed. Basal body proteins, which are transported to basal body from Golgi complex, are required in the formation of mature basal body. For example, two transition zone modules in *C. elegans*, NPHP module and MKS module, are required in generating the cilia base. The MKS module includes transmembrane lipid-interacting C2/B9 domain proteins MKS-1, MKSR-1, and MKSR-2, and also includes the MKS-3/TMEM67, MKS-5/RPGRIP1L, MKS-6/CC2D2A proteins. The NPHP module includes NPHP-1 and NPHP-4. Deleting one component from each module causes defective transition zone formation (Williams *et al*. 2011). In *C. elegans*, inversin compartment protein NPHP-2 (Inversin ortholog), enriched in middle segment, acts as a modifier of the MKS module and the NPHP module for transition zone formation in an amphid and phasmid sensilla-specific manner. Deleting NPHP-2 and one component of MKS module result in misplacement of transition zone, and short cilia length in amphid and phasmid sensilla (Warburton-Pitt *et al*. 2012). In *C. elegans*, the IFT-associated proteins, namely IFT proteins and BBS proteins, are anchored the basal body while forming the structure termed ciliary gate, which consists of transition zone and transition fibres (Williams *et al*. 2011). Then the axoneme is elongated by Intraflagellar transport (IFT) activities to form the middle and distal segment (Figure 1.5) (Inglis *et al*. 2007).

Intraflagellar transport (IFT) was first observed in *Chlamydomonas reinhardtii* flagella. They observed “rapid bidirectional granule-like particles movement along the length of *Chlamydomonas reinhardtii* flagella (Kozminski *et al*. 1993). The bidirectional movement is named anterograde movement and retrograde movement (Cole *et al*. 1998). Two motors kinesin-2 family (heterotrimeric kinensin-II) and homodimeric OSM-3, working cooperatively, in the anterograde direction (Cole *et al*. 1998). IFT particle sub-complexes A and B, large protein complexes that facilitate IFT cargo and motor connections, are bound to kniesin-II and OSM-3, respectively (Cole *et al*. 1998). BBS proteins stabilize the interaction between the IFT particles and the motors (Blacque *et al*. 2004; Snow *et al*. 2004; Ou *et al*. 2005). IFT dynein are bound to the IFT particles. IFT cargos/motors/particles/dynein complexes then move along the axoneme by kinesin-II and OSM-3 activities to form the middle segments. In the middle segment, the kinesin-II

movement slows down the fast OSM-3 movement. At the boundary of the middle segment, kinesin-II is released from IFT-complex. Only the OSM-3 motor moves the IFT cargos/particles/dynein to form the distal segment (Snow *et al*. 2004). For the retrograde direction, the redundant IFT motors, IFT particles, BBS proteins are moved back by IFT dynein activities to be recycled and used in next anterograde transport (Orozco *et al*. 1999). The discovery of Intraflagellar transport (IFT) was a milestone in cilia studies. After the discovery, many genes encoding IFT components are associated with a class of genetic diseases termed “ciliopathies”. For example, the gene encoding IFT protein IFT88/polaris is mutated in cystic kidney disease, one kind of ciliopathies (Pazour *et al*. 2002).

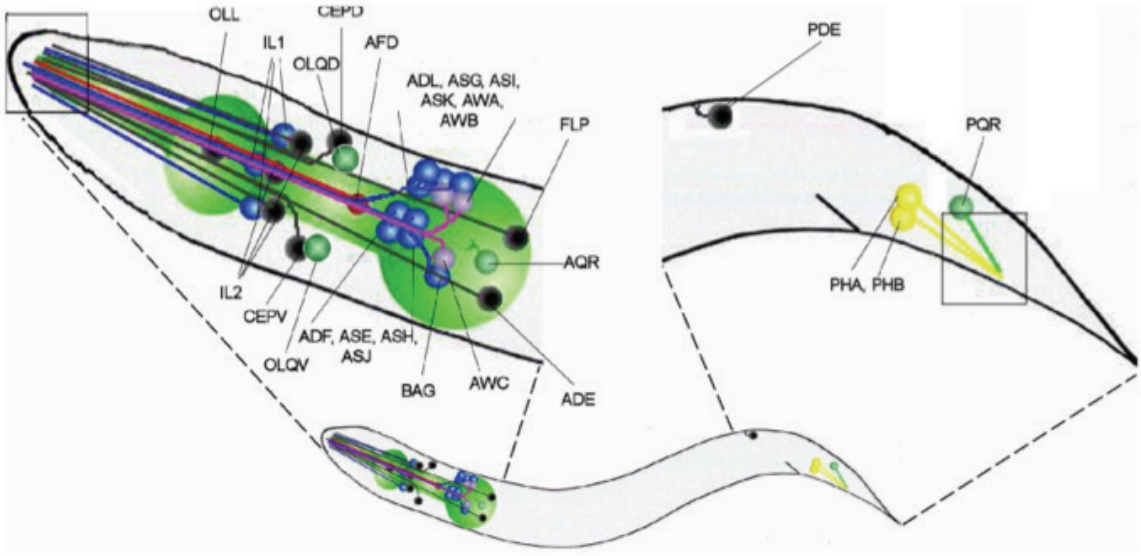


Figure 1.3. Location of ciliated sensory neurons in *C. elegans* hermaphrodites

Note. Figure shows all known ciliated sensory neurons cell bodies and associated dendrites relative positions in *C. elegans* hermaphrodites. Figure is adopted from (Inglis *et al.* 2007). This is a lateral view from left side. The neuron cell bodies are labelled with names. The names in the head region are explained as following. ASE, ASG, ASH, ASI, ASJ, ASK: amphids single rod. ADF, ADL: amphids pair rods. AWA, AWB, AWC: amphids wing. AFD: amphids finger. ADE: anterior deirid. IL1: inner labial 1. IL2: inner labial 2. OLL: outer labial. OLQD: outer labial quadrant dorsal. OLQV: outer labial quadrant ventral. CEPD: cephalic dorsal. CEPV: cephalic ventral. AQR: amphids pseudocoelomic. BAG: bags. FLP: flips. The names in the tail region are explained as following. PHA, PHB: phasmids A/B. PDE: posterior deirid. PQR: phasmids pseudocoelomic. The ciliated dendrites endings are indicated with black rectangle box.

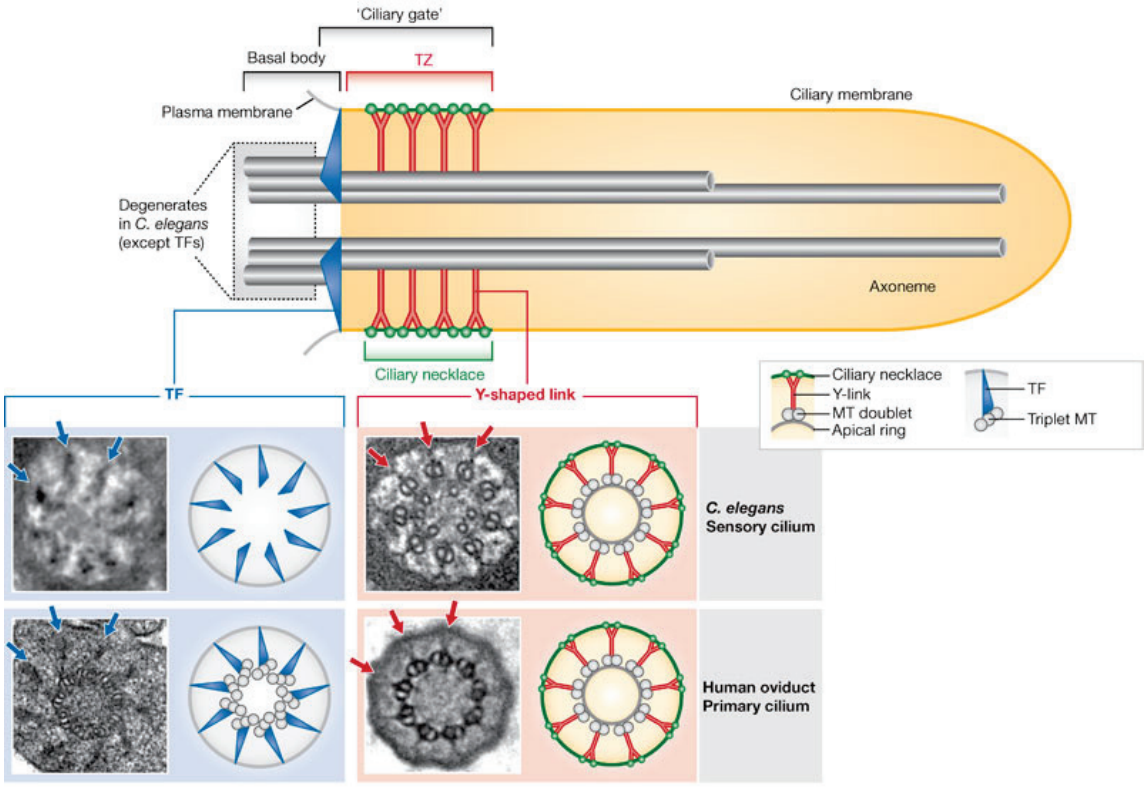


Figure 1.4. Structure of primary cilium base in *C. elegans*

Note. Fibre structures and TEM images in *C. elegans* cilium basal body and transitions zone. Figure is adopted from (Reiter et al. 2012). Cilium has a microtubule-based axoneme that is projected from a centriolar structure termed "basal body". Cilium basal body has "9x3+0" microtubules, surrounded by transition fibres (TF), shown in TEM. In *C. elegans*, microtubules are degenerated and only transition fibres composed of basal body, shown in TEM. The transition zone (TZ) has "9x2+0" microtubules. Y-shaped links in TZ make up the ciliary necklace on ciliary surface, shown in TEM and as beads on schematic.

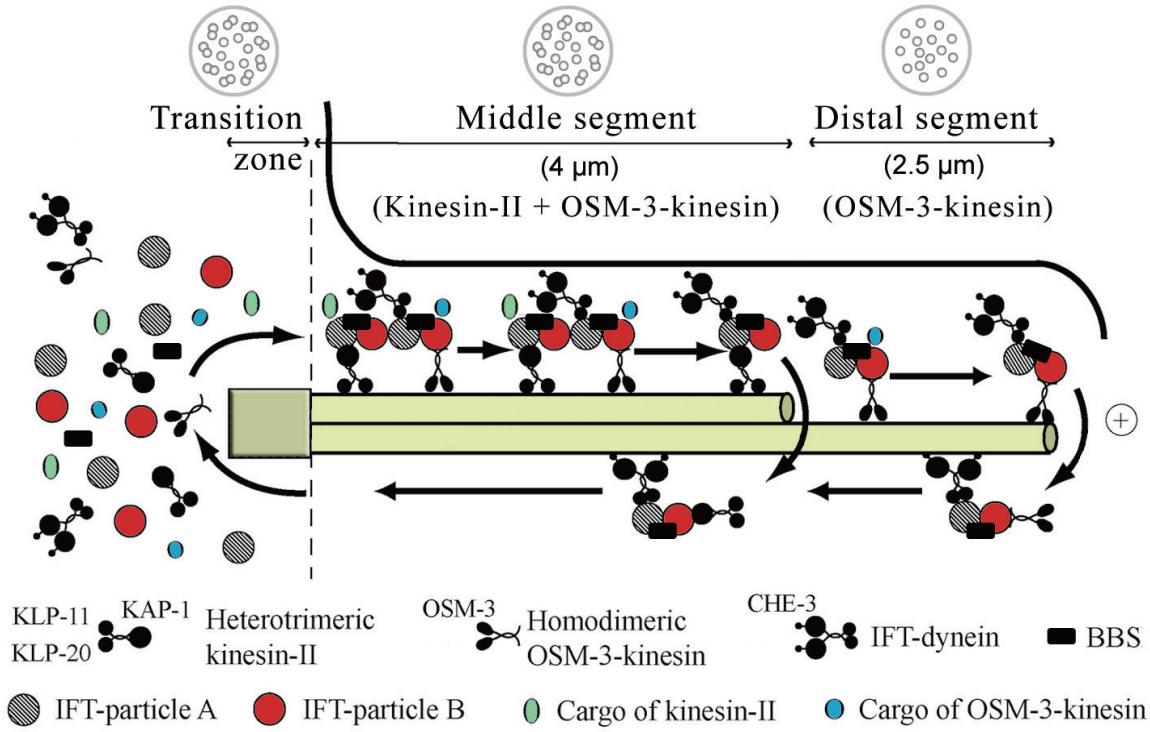


Figure 1.5. Intraflagellar transport (IFT) in *C. elegans*

Note. Ciliary components for IFT in extension of cilia are indicated in different shades. Figure is adopted from (Inglis *et al.* 2007). Components of IFT and ciliary cargos are assembled at transition zone (TZ). In the middle segment, kinesin-II and OSM-3-kinesin bind to IFT-particle A and IFT-particle B separately, and move along the axoneme in the anterograde (+) direction together with cargos and IF-dynein. In the distal segment, OSM-3-kinesin alone transports IFT-particles, cargos and IFT-dynein along the axoneme in the anterograde (+) direction. IFT dynein transports along the axoneme in retrograde (-) direction to recycle IFT-particles and kinesins to TZ. BBS proteins stabilize the association of kinesins and IFT-particles complex. The length of TZ (1 μm), middle segment (4 μm) and distal segment (2.5 μm) are shown with scale bars.

1.3. Ciliopathy and Meckel Gruber Syndrome (MKS)

Failure of ciliary gene functions causes severe genetic diseases. These diseases are classified into a type called ciliopathies. Ciliopathy patients usually have polycystic kidney disease, nephronophthisis, rentinitis pigmentosa, mental retardation, polydactyly, hepatic disease, bone malformation, diabetes, obesity and other problems with lungs, limbs, kidneys, brains and eyes (Hildebrandt *et al.* 2011; Davis and Katsanis 2012). Ciliopathy is defined as “genetically heterogeneous disorders that are caused by mutations in genes with products localized to cilium-centrosome complex” (Hildebrandt *et al.* 2011).

Ciliopathies are classified into mild, moderate and severe based on the onset symptoms. For example, Bardet-biedl Syndrome (BBS) is a mild ciliopathy, which causes developmental abnormalities and degenerative problems. Patients are often diagnosed in late childhood or early adulthood. Although there is no targeted treatment for BBS, clinical approaches can help patients (Dar 2001). An example of a moderate ciliopathy is Joubert Syndrome (JBTS). JBTS patients have mental retardation and ataxia. It is diagnosed in early childhood. Supportive clinical approaches are available to increase survival rate of JBTS patients (Maria *et al.* 1999; Karegowda *et al.* 2014). Severe ciliopathies always cause lethality at birth. Meckel Gruber Syndrome (MKS) is a severe ciliopathy.

Ciliopathies were genetically classified as dominant or recessive. Von Hippel–Lindau Disease (VHL) is an autosomal dominant ciliopathy caused by heterozygous germline inactivation of VHL gene (Kaelin 2003). BBS, JBTS and MKS are autosomal recessive disorders, with recessive mutations in multiple genes. More than 50 causal loci have been genetically identified in more than 15 clinically discrete ciliopathies (Badano *et al*. 2006; Baker and Beales 2009). In any of the clinically discrete ciliopathies, multiple causal genes have been found. For example, 16 genes are associated with over 80% clinically diagnosed BBS patients (Forsythe and Beales 2013). Among these causal genes, *BBS1* and *BBS10*, are the major pathogenically mutated genes. 23.1% patients have mutations in *BBS1* gene, and 20% of patients carry mutation in *BBS10* gene (Waters and Beales 1993-2003 (updated 2011)). Different causal loci in same gene or allelism at single locus result in different severity of clinically discrete disorders. For example, mutations in *NPHP6* gene cause JBTS, MKS and Nephronophthisis (NPHP) (Hildebrandt *et al.* 2011). MKS is severe and lethal while NPHP is mild and degenerative (den Hollander *et al.* 2006; Helou *et al.* 2007).

In my thesis, I am interested in gene *MKS1,* which is first identified as causal gene in Meckel Gruber Syndrome (MKS) patients. I have three reasons to study *MKS1* gene. First, MKS is an autosomal recessive lethal disorder. Meckel first defined this disorder in a report, as he found newborn patients of “occipital encephalocele, polydactyly, cleft palate, and large cystic kidneys” who died after birth (Meckel 1822; Salonen and Paavola 1998). The prevalence of MKS varied from 1 in 13,250 to 1 in 140,000 in different countries (Salonen and Norio 1984). Based on data from 1977 to

1982, two European countries have high prevalence: The prevalence is 1 in 3,000 in Belgium and 1 in 9,000 in Finland (Salonen and Norio 1984; Alexiev *et al*. 2006). MKS symptoms include polydactyly, polycystic kidneys, encephalocele, and hepatic defects, while some of the patients also have congenital heart abnormities and cerebellar hypoplasia (Gazioglu *et al.* 1998). MKS is such fatal disease that patients are born dead or die within few months after birth. No treatment is available now. Early diagnosis is applied to stop gestation of MKS babies with medical abortion (Gazioglu *et al.* 1998). MKS is usually diagnosed by ultrasound at 11^{th} to 14^{th} weeks of gestation (Pachi *et al*. 1989; Sepulveda *et al*. 1997; Mittermayer *et al*. 2004). Because multiple causal genes have been identified, it is possible to diagnose MKS based on DNA information (Roume *et al*. 1998; Kyttala *et al*. 2006; Smith *et al*. 2006; Baala *et al*. 2007; Delous *et al*. 2007; Tallila *et al*. 2008). For reproductive couples with a diagnosed MKS infant, the chance of having a MKS baby in subsequent pregnancy is 25% each time (Alexiev *et al.* 2006).

As of 2009, five causal genes, *MKS1, TMEM67, CEP290, RPGRIP1L* and *CC2D2A,* have been identified in MKS (Roume *et al*. 1998; Kyttala *et al*. 2006; Smith *et al*. 2006; Baala *et al*. 2007; Delous *et al*. 2007; Tallila *et al*. 2008). Up to 2012, 17 causal genes (*NPHP3, TMEM216, TCTN2, BBS2, BBS4, BBS6, BBS10, B9D1, B9D2, KIF7, WDPCP, TTC218* and the five previously identified genes) were identified as casual loci in MKS patients (Karmous-Benailly *et al*. 2005; Stoetzel *et al*. 2006; Bergmann *et al*. 2008; Leitch *et al*. 2008; Kim *et al*. 2010; Riazuddin *et al*. 2010; Davis *et al*. 2011; Dowdle *et al*. 2011; Hopp *et al*. 2011; Sang *et al*. 2011). A recent study in Arabian MKS families identified three new causal genes *C5orf42, EVC2* and *SEC8* (Shaheen *et al.* 2013). Causal genes in MKS are allelic with those in JBTS, NPHP, BBS, Leber congenital amaurosis (LCA) and Senior-Loken syndrome (SLS). For example, these genes are causal in different ciliopathies: *MKS1* is causal for MKS and BBS, *TMEM67* for MKS and JBTS, *CEP290* for MKS, NPHP, BBS, LCA and SLS, *RPGRIP1L* for MKS, JBTS, NPHP and SLS, *CC2D2A* for MKS and JBTS (Roume *et al*. 1998; Kyttala *et al*. 2006; Smith *et al*. 2006; Baala *et al*. 2007; Delous *et al*. 2007; Tallila *et al*. 2008). Among 20 identified causal genes in MKS, *MKS1* and *TMEM67 (MKS3)* are the two major causal genes, with majority of patients carrying mutations in these two genes. For example, Majority of Finnish MKS families carried *MKS1* mutations. *TMEM67 (MKS3)* mutations were common in Asian MKS families. *MKS1* and *TMEM67 (MKS3)* are also

major causal genes in non-consanguineous European origin populations (Consugar *et al.* 2007).

Second, *MKS1* is mutated in majority of MKS patients, and is also mutated in BBS patients (Kyttala *et al*. 2006; Leitch *et al*. 2008). In 1995, mutations in a group of MKS patients were mapped to Chromosome 17q21-q24 in the 13cM region (Paavola *et al.* 1995). The first reported locus resides in the *MKS1* gene. In 2005, studies on suspected chromosome regions of 40 Finnish families and 24 non-Finnish families confirmed that 26 Finish families and 3 non-Finish Families have the heterozygous *MKS1* mutations in parents and homozygous *MKS1* mutations in affected children (Kyttala *et al.* 2006). The mutations were large intronic deletion, large exonic deletion, small exonic insertion or small intronic transition. *MKS1* encodes a B9 domain protein (Tammachote *et al.* 2009). C2/B9 domain binds to Ca^{+}/lipids, and B9 domain proteins usually participate in vesicle trafficking and membrane fusion (Nalefski and Falke 1996). MKS1 protein is localized to basal body of cilia (Dawe *et al*. 2007). MKS1 is important in cilia formation and epithelial morphogenesis (Dawe *et al.* 2007; Tammachote *et al.* 2009; Cui *et al.* 2011; Warburton-Pitt *et al.* 2012). Studies in mouse have shown that MKS1 protein, through interaction with MKS3/Meckelin, is important to translocation of centriole to cell plasma membrane during basal body formation in most, but not all tissues (Dawe *et al*. 2007; Weatherbee *et al*. 2009; Cui *et al*. 2011). Also MKS1 protein interacts with Shh signaling components and Wnt signaling components in mouse models (Dawe *et al.* 2007; Wheway *et al.* 2013). The *Mks1* null-mutant mice have abnormal limb bud formation and defects in kidney, nerve code and cochlea (Cui *et al.* 2011).

Third, gene *xbx-7/mks-1* is identified as ortholog of *MKS1* in *C. elegans* by reciprocal BLASTp of the MKS1 polypeptide (Kyttala *et al.* 2006). The gene *mks-1* expresses in a subset of ciliated neurons in *C. elegans* (Efimenko *et al.* 2005). The gene product MKS-1 is localized to transition zone in *C. elegans* (Williams *et al.* 2008). Studies have focused on the role of MKS-1 in transition zone and basal body proteins complex of cilia. MKS-1, MKSR-1 (B9D1) and MKSR-2 (B9D2), three B9 domains proteins in *C. elegans*, genetically interacted with each other (Williams *et al*. 2008; Bialas *et al*. 2009). Further studies in *C. elegans* demonstrated that the three B9 domain proteins, MKS-1, MKSR-1, MKSR-2, form a complex with MKS-3, MKS-5, MKS-6 to recruit mother centriole docking to membrane. The MKS complex works cooperatively

with NPHP (nephronophthisis) complex, which includes NPHP-1 and NPHP-4. Deleting components of both complexes cause abnormal cilia in *C. elegans*. Thus MKS/NPHP proteins recruit the basal body, and form ciliary gate to block non-ciliary proteins in cilia formation (Williams *et al.* 2008; Bialas *et al.* 2009; Williams *et al.* 2011; Warburton-Pitt *et al.* 2012). 45 different mutant alleles of *mks-1* are available in *C. elegans*. 43 of them were generated by "the million mutation project", and further analysis on these *mks-1* mutants phenotype is required (Thompson *et al.* 2013). The other two alleles are *tm2705* (large deletion in exons and intron) and *otn392* (single substitution in intron). *otn392* is carried by a strain with multiple mutations obtained from EMS mutagenesis, and is detected by whole genome sequencing (Sarin *et al*. 2010). *tm2705* deletion, a 190-bp deletion of *mks-1* that begins in intron 2 and includes partial exon 4, cause the splicing of exon 2, and thus the protein product lacks 70 amino acids near N-terminus. *tm2705* is a hypomorphic allele, the splicing does not abrogate the B9 domain (Williams *et al*. 2008). Single *tm2705* mutant exhibit normal Intraflagellar transport (Bialas *et al*. 2009). Single *tm2705* mutant does not have dye-filling defect phenotype (non-Dyf), but double mutants that carry tm2705 allele and a mutant allele of *nphp-1* or *nphp-4*, which both belong to NPHP complex that participate in basal body formation, have a severe Dyf phenotype (Williams *et al*. 2008; Warburton-Pitt *et al*. 2012). In accordance to Dyf phenotype, the dendrites of phasmid neurons are not properly extended in *mks-1 (tm2705)* and *nphp-4* double mutants (Williams *et al*. 2008).

1.4. Transcriptional Regulation in *C. elegans*

Although there are many studies on MKS-1 functions, there are limited studies on transcriptional regulation of gene *mks-1*. Thus, I am interested in studying transcriptional regulation mechanism of *mks-1*.

In eukaryotes, RNA polymerase II catalyze the transcription of DNA to mRNA. To initiate the transcription, RNA polymerase II must be recruited to the transcription initiation site. To recruit RNA polymerase II, interactions between transcription factors and cis-regulatory sequences are required. The process is called transcription regulation (Lemon and Tjian 2000). The first step is to assemble a pre-initiation complex (PIC), which consists of RNA Pol II, a set of general transcription factors (GTFs) that regulates

all or most gene expression, and mediator proteins that directly interact with specific transcription factors (Lemon and Tjian 2000; Orphanides and Reinberg 2002). Recruitment of PIC is regulated by activities of transcription factors and histone modifications as well (Chan and La Thangue 2001). After PIC is recruited to core promoter regions, GTFs activities enable RNA pol II binding to specific transcription start site (i.e. start codon). The core promoter region contains conserved cis-regulatory sequences that interact with GTFs. In eukaryotes, TATA-box Binding Protein (TBP) interacts with a TA rich sequence (TATA box), which is an important activity for recruiting RNA Pol II binding (Lichtsteiner and Tjian 1993). RNA Pol II must be phosphorylated on an extended C-terminal domain (CTD) to an active form to start transcription (Seydoux and Dunn 1997). The phosphorylation is assisted by activates of cyclin-dependent kinase on some GTFs (Bentley 2005).

In most cases, cis-regulatory elements are located in the intergenic region from the 5'UTR of the interested gene to the nearest upstream gene. These regions include core promoter, proximal promoter and distal promoter. Core promoter is the minimal promoter sequences that contain a set of conserved cis-regulatory elements to initiate transcription. Core promoter is the region where GTFs bind, where RNA Pol II binds and where transcription start site locates. Grishkevich *et al.* studied the conserved cis-regulatory elements among different core promoter regions in *C. elegans*. They identified five conserved elements, including Sp1 like site for general transcription factor SP1 protein binding, T-blocks that correlate to nucleosome eviction and gene expression level, TATA box, trans-splicing site and Kozak sequence that include the transcription start site (Grishkevich *et al*. 2011). Proximal promoter region is located close to the target genes. In *C. elegans* this region is usually located within 2 kb upstream of the translation start codon (Okkema and Krause 2005; Reinke *et al*. 2013). Most specific transcription factors binding sites reside in proximal promoter regions. The modENCODE project has set up a series of database of transcription factors and cis-regulatory sequences information in *C. elegans*. In a modENCODE study to identify cis-regulatory elements for transcription factors, majority of the binding sites lie within 500 bp of the transcription start site near gene coding region (Niu *et al*. 2011). Distal promoter region is located far upstream of the coding sequences. For example, in *C. elegans*, the Hox gene *lin-39* expression requires regulation by cis-elements ~30 kb

upstream of the coding region (Wagmaister *et al*. 2006). Regulation in this region usually causes a weaker influence on gene expression level (Kroeger *et al*. 1987).

The natural promoter, where all cis-elements are found, covers more than merely upstream of coding regions. Cis-elements are found in downstream and large introns of targets genes as well. For example, in the same study of regulatory sequences for Hox gene *lin-39* expression in *C. elegans*, cis-elements are also found ~30 kb downstream of coding region (Wagmaister *et al*. 2006). In *C. elegans* most introns are smaller than 100 bp. Small introns less 100 bp are not likely to contain cis-regulatory elements. However, intronic regulatory sequences have been identified in larger introns. For example, Reid *et al.* identified two intronic regulatory elements that regulate expression of hypoxanthine phosphoribosyltransferase (*HPRT)* gene in mouse embryonic stem cells (Reid *et al*. 1990).

Transcription factors, also called trans-acting factors, are proteins that bind to cis-regulatory elements to regulate gene expression. There are two types of transcription factors, general and specific. General transcription factors (GTFs), namely TFII1, TFIIB, TFIID, TFIIE, TFIIF and TFIIH, recruit RNA Pol II to transcription start site directly (Matsui *et al*. 1980; Orphanides *et al*. 1996). They are required for most gene expressions. Specific transcription factors regulate gene expression in a tissue-specific, cell-specific, gene-specific or pathway-specific manner (Britton *et al*. 1998; Chang *et al*. 2003). Specific transcription factors either activate (by activator activities) or repress (by repressor activities) target gene expression. Non-covalent interactions between transcriptional co-factors and DNA-binding transcription factors also regulate target gene expression (Shi and Mello 1998).

Noncoding microRNA specifically regulate gene expression in post-transcriptional level as well. Mature microRNA binds to miRNA target sites in mRNA transcript, and thus repress translation of mRNA (Engels and Hutvagner 2006). Mature microRNA is short single strand RNA that is cleaved from hairpin-structure pre-miRNA by enzyme dicer. Pre-miRNA is formed from a pri-mRNA transcript that is transcribed from DNA by RNA Pol II (Lee *et al*. 2004). MicroRNA was first discovered in *C. elegans* (Lee *et al*. 1993), and many *C. elegans* microRNAs were identified in the modENCODE project (Kato *et al*. 2009).

In chromatin level, histone modification regulates gene expression because DNA is wound into nucleosomes that are composed of histone protein complexes (H1, H2A, H2B, H3 and H4). Residues on amino terminal of histones are modified by phosphorylation, methylation, acetylation and ubiquitination (Bannister and Kouzarides 2011). Such modifications usually have effect on gene transcription by changing chromatin density to facilitate or repress association of transcription factor and DNA, or by proving specific recruitment sites for transcription factors (Stimpson and Sullivan 2010). Studies of histone modification during *C. elegans* development have identified several conserved chromatin complexes that guided the modifications, and thus regulating target gene expression. For example, NuRD (Nucleosome Remodeling and Deacetylase) complex deacetylates lysine 9 of histone H3 (Brehm *et al*. 1998), and then histone methyltransferases (HMT) could methylate the residue to suppress target gene expression (Andersen and Horvitz 2007).

C. elegans is an ideal organism to study transcription regulation, because its transparent body is ideal for monitoring gene expression in live worms by using fluorescent proteins, and because its small genome size is ideal for searching cis-regulatory elements by using Bioinformatics tools. Computational search identified 934 transcription factor-encoding genes by using Gene-Ontology terms and DNA binding domain sequences in the worm genome, which composes of ~ 4.6% of all protein coding genes in *C. elegans* (Reece-Hoyes *et al*. 2005). Genome- wide binding profiles of 120 transcription factors were studied in the modENCODE project, which aims to explore transcriptional regulation elements in model organisms (Gerstein *et al*. 2010).

There are many ways to study transcriptional regulation in *C. elegans*. Applying mutagenesis and genetic screens can identify mutations in transcription factors (Singh and Han 1995). RNA interference (RNAi) disrupts the regulation pathways and helps analyze functions of regulators (Powell-Coffman *et al*. 1996). The presence of mRNA could be visualized by *in situ* hybridization (Seydoux and Fire 1994). Chromatin immunoprecipitation (ChIP) is used to directly identify binding sites for transcription factors *in vivo* (Oh *et al*. 2006). Many transcription factors that regulate different development pathways including pharynx development, vulva development and neuron development have been identified through genetic screens. For example, *pha-4*, the

central regulator in pharynx development, was identified in a genetic screen for larva lethal phenotypes (Mango *et al*. 1994).

Two conserved transcription factor families are identified in ciliary gene regulation: RFX (Regulatory Factor X), discovered in both motile and primary cilia, and FOX (forkhead protein) family, only discovered in motile cilia (Reith *et al*. 1990; Clevidence *et al*. 1993).

The first RFX was identified in a search of proteins biding to a X-box motif DNA sequence in MHC class II gene promoter regions (Reith *et al*. 1990). X-box motif was identified by searching for conserved sequence in 5' flanking regions of MHC class II gene molecules (Saito *et al*. 1983). RFX family has a characterized winged-helix DNA-binding Domain (DBD). RFX use this DBD to bind to the minor groove of DNA in X-box (Gajiwala *et al*. 2000). RFX are highly conserved through species. Nine RFXs have been identified in human, including RFX1-9 (Emery *et al.* 1996; Chu *et al.* 2010). In *Drosophila*, RFX1 and RFX2 are identified. CRT1 is identified in *S. cerevisiae*. SAK1 is identified in *S. pombe* (Huang *et al.* 1998; Durand *et al.* 2000; Thomas *et al.* 2010). DAF-19 is the sole known RFX member in *C. elegans* (Swoboda *et al.* 2000). A subgroup of RFX, including *C. elegans* DAF-19, *Drosophila* RFX, human RFX1-4, RFX6 and RFX8, are connected to transcriptional regulation in cilia (Chu *et al*. 2010). This subgroup of RFX shared five conserved domains including an activation domain, DBD, domain B, domain C and dimerization domain, which are conserved structurally and functionally (Choksi *et al.* 2014).

DAF-19 is the sole known member of RFX in *C. elegans* (Swoboda *et al*. 2000). DAF-19C isoform expresses specifically in all ciliated neurons in *C. elegans* (Senti and Swoboda 2008). *daf-19* null-mutants have defective cilia, and have phenotypes of constitutively dauer formation and Dyf phenotype (Dye-filling defects) (Swoboda *et al*. 2000). X-box is a predictor for ciliary genes (Chen *et al.* 2006). Up to now, the featured X-box motif, for DAF-19 characteristic DNA domain binding, is detected upstream of all known ciliary gene in *C. elegans*. Thus, Efimenko *et al.* use bioinformatics tolls to predict ~ 750 *xbx* genes (genes with X-box motif) in *C. elegans* by the search of X-box motif in gene promoter regions (Efimenko *et al.* 2005). They generated transcriptional fusion constructs of 27 predicted *xbx* genes and injected the constructs into *daf-19* null-mutant

to observe transgene expression. They validated that expression of 22 *xbx* genes (out of 27) are DAF-19 dependent, based on transgene expression in wildtype and *daf-19* null-mutant. Efimenko *et al.* divided DAF-19 dependent genes into two groups. The first group expresses in all ciliated neurons. It is predicted that genes in first group are required by cilia formation in most ciliated neurons and are strongly regulated by DAF-19. The second group expresses in subset of ciliated neurons. It is predicted that genes in second group are required by cilia formation in specific ciliated neurons and are partially regulated by DAF-19. Examples of RFX-targeted ciliary genes and their correlations to cilia components are shown in Figure 1.6 (Thomas *et al*. 2010). The gene *mks-1* was detected in second group of ciliary genes in Efimenko *et al*. studies (Efimenko *et al*. 2005). The X-box sequence for *mks-1* gene is GTCACCATAGGAAC. The expression of transgene *pmks-1::gfp* in *daf-19* null-mutant was strongly diminished in Efimenko *et al*. study, while it is absent in Williams *et al*. study (Efimenko *et al*. 2005; Williams *et al*. 2008). A second motif is identified in DAF-19 dependent ciliary gene promoter regions recently. The C-box, an 8-to-11bp motif, function as an enhancer in DAF-19 dependent gene expression, although transcription factors that bind to C-box and binding domains are yet to be known (Burghoorn *et al.* 2012).

Forkhead box protein J1 (FOXJ1) was first cloned from rat lung cDNA library in PCR for forkhead domain sequences (Clevidence *et al*. 1993). FOXJ1 belongs to FOX family, which has a forkhead domain. It is identified that FOX family regulates ciliary gene expression in motile cilia of mouse, *Xenopus*, zebrafish and *Drosophila* (Chen *et al*. 1998; Brody *et al*. 2000; Stubbs *et al*. 2008; Yu *et al*. 2008; Newton *et al*. 2012). FOXJ1 knock-out mouse had complete loss of axonemes in cilia (Chen *et al.* 1998; Brody *et al.* 2000). Fox transcriptional factor in *Drosophila*, FD3F, regulates cilia formation in auditory and chordotonal neurons (Cachero *et al.* 2011; Newton *et al.* 2012). Because *C.elegans* does not have motile cilia, no ortholog of FOX family is identified in *C. elegans* (Mazet *et al*. 2003).

Besides RFX family and Fox family, cell-specific regulators in cilia formation have been identified by multiple studies in different organism. In mouse, for example, CRX (cone-rod homeobox containing gene) regulates ciliated photoreceptor cells differentiation (Freund *et al.* 1997; Furukawa *et al.* 1997). Another homeobox transcription factor, NOTO (notochord homeobox), regulates motile cilia formation in

ventral node of mouse (Beckers *et al*. 2007). In *Drosophila*, ATO (Atonal) regulates motile cilia formation indirectly in chordtonal neurons to activate expression of *fd3F* and *Rfx* (Cachero *et al*. 2011). In *Xenopus* epidermis and *mouse* airway, multicilin, a coiled-coin domain containing protein, is required for multi-ciliated cell formation (Stubbs *et al*. 2012). In *C. elegans*, transcription factor FKH-2 (forkhead 2) specifically works downstream of DAF-19 to regulate kinesin-II subunit gene *kap-1* expression in olfactory neuron AWB (amphid wing B) (Mukhopadhyay *et al*. 2007b). The miRNA targeting ciliary genes in post-transcriptional level is also identified in mouse. For example in ciliated cells in mammalian inner ears, miR-96 targets *oldf-2* gene, which encodes a component of cilia basal body (Mencia *et al*. 2009). These data suggested a complicated regulating system of ciliary genes. In specific cells, other transcription factors regulate downstream, upstream or as co-factor of RFX (in primary cilia formation) or FOXJ1 (in motile cilia formation) for some ciliary gene expression. And in post-transcription level, miRNA regulates some gene expression in specific cells.

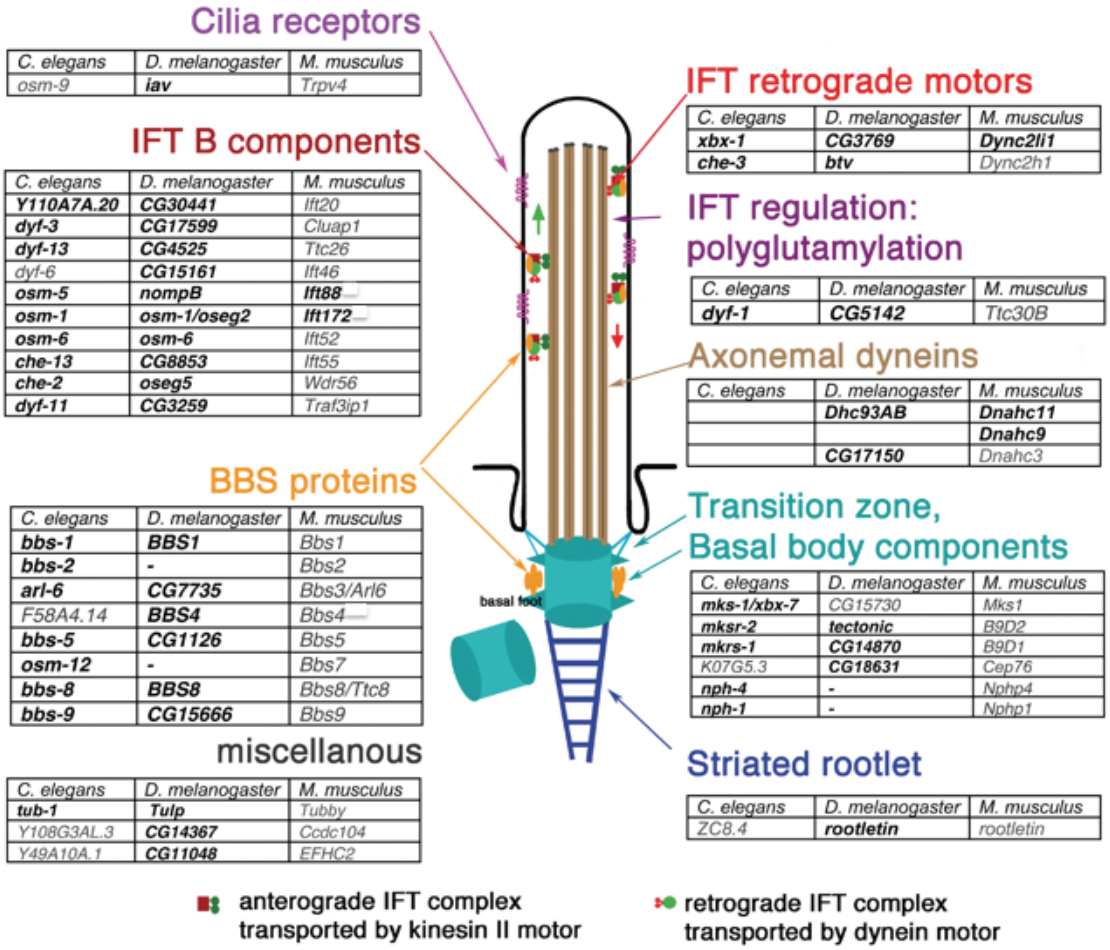


Figure 1.6. RFX targeted ciliary genes

Note. Gene orthlogs related to different components of cilia in *C. elegans, Drosophila* and mouse. Figure is adopted from (Thomas *et al.* 2010). Simple schematic of a primary cilium basal body, middle segment and distal segment is shown. Gene orthologs, of which protein products correlate to different ciliary components, are listed in tables. Genes in bold are found to be regulated by Regulatory Factor X (RFX). Other genes were not tested.

1.5. Applying Genetic Screening to identify transcription factors

Among methods to study transcriptional regulation, forward genetic screening provides a direct way to identify unknown transcription factors of interested genes. In forward genetic screens, non-directed genome-wide mutagenesis is applied to introduce random mutations. One way of mutagenesis is to treat wildtype animals with chemical mutagens such as ethyl methanesulfonate (EMS) (Brenner 1974). Mutants defective in

certain biological processes are picked up during screen. The genetic mapping, whole genome sequencing and molecular analysis are used to identify unknown mutated gene(s) after genetic screening. In my thesis study, I applied genetic screens to find mutants defective in *mks-1* expression.

Sydney Brenner first introduced genetic screening method to the nematode *C. elegans*. In Brenner's work, he treated wildtype *C. elegans* with ethyl methanesulfonate (EMS), an alkylating mutagen, to introduce mutations (Brenner 1974). During the screening, he generated the mutant population in two ways. 30-40 plates with one mother worm each were prepared. In M set, The EMS treated mother worms were removed after it laid around 50 F1 progeny. The F2 in M set were screened, and only one mutant was picked up from each plate. In S set, mother worms were not removed. 5 F1 from same mother were transferred to new plates, with one worm per plates. The F2 were screened and one mutant was picked up from the same F1. The S set avoided the bias of slow growing mutants and mutant phenotypes that only expressed in adult form. Mutant worms were different in size, shape and movement. The screening successfully identified mutants of uncoordinated (*unc* genes), roller (*rol* genes), dumpy (*dpy* genes), small (*sma* genes), long (*lon* genes), blistered and abnormal formation. Mutants were crossed to wildtype N2 males to determine whether the mutations are recessive, dominant or semi-dominant. Meanwhile, one mutant is crossed to another mutant in a complementation test. If F1 progeny are wildtype, two mutants have different mutations. If F1 progeny still possesses mutant phenotype, two mutants have the same mutation. Only mutants that generate reproductive males can be tested in a complementation test. Mutations are mapped using phenotypic gene markers. Mutant of interest is crossed to the mutant with phenotypic markers (usually recessive mutations in marker genes on different chromosomal locations). The F1 progeny are double heterozygous with wildtype phenotypes. F1 progeny are self-crossed to generate F2 progeny. F2 progeny has segregation of two mutant phenotypes. If mutant of interest is recessive, double homozygous F2 generation segregate have double mutant phenotypes. Genetic distance is the recombination rate between two loci (in every 100 chromosomes). Recombination rate (genetic distance) is calculated with the equation: genetic distance = $\sqrt{}$ (percentage of double homozygous in whole F2 generation*100) (cM). About 100 genes were identified in Brenner's genetic screens, of which results published in 1974.

After Brenner introduced genetic screening methods to *C. elegans*, it was improved to find genes in signalling pathway, modifiers (enhancers or suppressors), lethal genes, maternal-effect genes, or drug-resistant genes. Many transcription factors have been identified through genetic screens. For example, Mango *et al*. applied forward genetic screens to isolate mutants of first-stage larva arrest. In their screens, they identified *pha-4* gene, which encodes a transcription factor in pharynx development (Mango *et al*. 1994; Kiefer *et al*. 2007). In studies of neurogenesis, genetic screens are widely applied to identify genes that control neuronal specification. In screens for behavioural mutants, transcription factors that regulate specific neuron differentiation were identified. For example, *che-1* mutant is isolated in forward genetic screens for mutants with defects in chemotaxis behaviour (Dusenbery *et al*. 1975; Chang *et al*. 2003; Uchida *et al*. 2003). The gene *che-1* encodes a zinc finger transcription factor CHE-1 that specifically regulates ASE neuron differentiation (Sarin *et al*. 2009). Transcription factors that regulate axon positioning or neuron migration are also identified from genetic screens for mutants that are defective in axon positioning and neuron migration. For example, *ceh-17* mutants were obtained in reverse genetic screens to reveal *ceh-17* role in regulating axon growth and navigation (Pujol *et al*. 2000).

Some transcription factor mutants do not have visible phenotypes. Markers as green fluorescence proteins (GFP) were introduced to worms in genetic screens to identify transcription regulators. Doitsudou *et al*. provided a good example of using GFP marker in genetic screens (Doitsidou *et al*. 2008). In a forward genetic screen to isolate mutants defective in dopaminergic cell fate, they used promoter of dopaminergic-neuron specific gene fused to GFP transgene to track dopaminergic neurons. They use EMS to mutagenize transgenic worms. They screened for mutants with reduced GFP expression in one or multiple dopaminergic neurons.

1.6. Aims of the thesis

Many studies have been carried out to functionally analyse *MKS1* orthologs in mouse, *Drosophila* and *C. elegans*. I studied *mks-1* transcription regulation for two reasons. First, *MKS1*, human and mouse ortholog of *C.elegans mks-1* is the first

identified casual loci in Meckle Syndrome, one type of ciliopathies. Ciliopathies are caused by genetic defects in cilia formation, and thus mutations in same causal gene could lead to different syndromes. Mutations in *MKS1* also cause other ciliopathy, such as Bardet Biedl Syndrome (Kyttala *et al*. 2006; Leitch *et al*. 2008). Ciliopathies could cause fatal symptoms that lead to early death. No effective treatment is available to cure ciliopathies. Although each ciliopathy has low prevalence, the prevalence of overall ciliopathies is high because more than 100 ciliopathies have been identified. The prevalence of overall ciliopathies is 1 in 1000, which is as high as common genetic defects Down Syndrome (Hook 1983). Thus, it is important to understand the molecular mechanism that cause ciliopathies. Second, transcription regulation plays an important role in generating proper proteins. Some identified causal loci of ciliopathies locate in genes that encode transcription regulator. One causal locus of Meckel Syndrom resides in gene *KIF7* (Putoux *et al*. 2011). *KIF7* (kinesin family member 7) encodes a cilia-associated protein that regulates GLI1 and GLI2 transcription factor through sonic hedgehog (SHH) pathway (Dafinger *et al*. 2011). It is important to understand the regulation mechanism of ciliary proteins.

However, there were only few studies about transcriptional regulation of *MKS1*. In *C. elegans*, the only known transcriptional factor of *mks-1* is DAF-19. DAF-19 is the sole transcriptional factor of the RFX family in *C. elegans*. Multiple lines of evidence reveal that *mks-1* regulation involves a group of transcription factor including DAF-19. First *daf-19c* isoform expresses in all ciliated neurons but *mks-1* expresses in subset of ciliated neurons (Efimenko *et al*. 2005). A mechanism that aid DAF-19 to turn on *mks-1* expression in specific cell is yet to be known. A specific transcription factor is likely to be involved in this process to turn on *mks-1* expression in subset of ciliated neurons. Second, other regulatory sequences have been discovered in promoter region of ciliary gene including *mks-1* for DAF-19 dependent expression. C-box was discovered as an enhancer in the promoter of *daf-19* dependent ciliary genes that already contains RFX characteristic cis-element X-box (Burghoorn *et al*. 2012). It is not known what kind of transcription factor interacts with the C-box in regulating *daf-19* dependent ciliary gene expression. Third, many cell-specific transcription factors have been identified in regulating ciiliary gene expression in model organisms including *C. elegans* e.g. (Furukawa *et al*. 1997; Beckers *et al*. 2007; Mukhopadhyay *et al*. 2007b; Cachero *et al*.

2011; Stubbs *et al*. 2012). Co-regulation by RFX3 and FOXJ1 (i.e. FD3F in *Drosophila*) exists in regulating ciliary genes in motile cilia formation in mouse and *Drosophila* (Newton *et al*. 2012; Didon *et al*. 2013). Thus, I hypothesize in my thesis that *mks-1* is regulated by other transcriptional factors together with DAF-19.

In my thesis, I have two specific aims. First, I would like to generate transgenic worms that stably express GFP driven by *mks-1* promoter. I would like to examine expression patterns of the reporter transgene, so I could use this strain for genetic screens. Second, I would like to isolate mutants in candidate transcription factor(s) of *mks-1*.

In Chapter 2, I describe generating the transgenic worms that stably express GFP driven by *mks-1* promoter. Moreover, I observed *pmks-1::gfp* expression in *daf-19* mutant.

In Chapter 3, I mainly discuss the technical details of my two rounds of genetic screens, characterizations of mutants, and preliminary genetic analyses. I successfully isolated a X-linked recessive mutant with changed *pmks-1::gfp* expression pattern.

In Chapter 4, I discuss the value of this thesis work, and improvements for methods. I also propose the future work.

The appendix is about an alternative approach that I applied: candidate gene approach. Although I did not find significant targets here, results may be valuable for other researchers.

Chapter 2.

Generation of a transgenic strain that stably expresses GFP driven by *mks-1* promoter

2.1. Background

The ciliary gene *mks-1* expresses in amphids and inner labial neurons in the head and phasmids in the tail (Figure 2.1A and 2.1B) (Efimenko *et al.* 2005; Williams *et al.* 2008). Further observation at subcellular level resolution demonstrated that MKS-1 protein concentrates to basal body/ transition zone in cilia (Figure 2.1C) (Bialas *et al.* 2009).

The expression of *mks-1* is *daf-19* dependent (Efimenko *et al.* 2005). DAF-19 is the sole RFX family in *C. elegans* that regulates all known ciliary gene expression (Swoboda *et al.* 2000). The gene *daf-19* encodes three isoforms. DAF-19^{C} isoform is expressed in all ciliated sensory neurons (Figure 2.2) (Swoboda et al. 2000; Senti and Swoboda 2008). In my study, I used *daf-19* mutant allele *m86*. *daf-19 (m86)* mutant is generated by EMS mutagenesis, and selected from a genetic screening for mutants of constitutively dauer larva formation (Malone and Thomas 1994). The *m86* allele carries a nonsense mutation C ➔ T, resulting in an Arginine change to a premature stop codon before the DNA binding Domain (Swoboda *et al*. 2000). Thus, *m86* is a null-allele of *daf-19*. Electron Microscopy observation of *daf-19 (m86)* mutants showed that ciliated endings of sensory neurons are entirely missing (Perkins *et al*. 1986). In accordance to this observation, the *daf-19 (m86)* mutants have Dyf phenotype, and constitutively form dauer larva stage (67% at 15°C and 85% at 20°C) (Swoboda *et al*. 2000). Efimenko *et al.* generated transgene GFP constructs by inserting 2.1 kb promoter region of *mks-1* into GFP expression vector pPD95.77. They microinjected the constructs into *daf-19 (m86)* mutants and wildtype worms, together with co-injection markers. They observed

diminished *mks-1* expression in *daf-19 (m86)* mutant (Efimenko *et al*. 2005). Williams *et al*. also generated a transgenic strain by injecting transcriptional fusion constructs *xbx-7::YFP* to *daf-19 (m86)* mutants (*xbx-7* is the name used for *mks-1* in their study). They observed abolished *xbx-7::YFP* transgene expression in *daf-19(m86)* mutant background (Figure 2.1D). They then crossed the transgenic strain to *daf-19* wildtype worms, and the transgene expression is restored after the outcross (Williams *et al*. 2008). X-box is the putative target site for DAF-19. It is a 14bp to 20bp short sequence (Chu *et al.* 2012). In both Efimenko *et al.* (Efimenko *et al.* 2005) and Jeffery Chu's predictions (Dr. J. Chu, personal communication), they found X-box sequences in *mks-1* promoter region. Thus, it is confirmed that *mks-1* is regulated by DAF-19.

In my study, I need a transgenic strain that stably expresses *mks-1* promoter driven GFP in a wildtype background for EMS mutagenesis and genetic screening. Previously, Efimekno *et al.* generated a transgenic strain by co-injecting the *rol-6 (su1006)* dominant marker and GFP expression vector that carries *mks-1* promoter sequences into wildtype worms. The transgenic strain OE3152 *ofEx120 [R148.1::gfp;rol-6(su1006)*] has *pmks-1::gfp* expression in subset of ciliated neurons, including labial neurons, amphids and phasmids (Figure 2.1A). However, the OE3152 transgenic strain that Efimenk *et al.* generated has a Rol phenotype. Thus, it is hard to use for genetic screening. I decided to microinject the transcriptional fusion construct of *pmks-1::gfp* into *dpy-5 (e907)* recessive mutant, with *dpy-5* rescue plasmids. The goal in is to generate a transgenic *C. elegans* strain that has a wildtype phenotype with stable expression of GFP constructs driven by *mks-1* promoter. With help from my colleagues, I generated the promoter driven transcriptional expression of *mks-1*, using the same *mks-1* promoter sequences that Efimenko *et al.* used (Efimenko *et al.* 2005). The 2.1 kb promoter sequence covered the upstream intergenic region between *mks-1* and *R148.2*. X-box is located inside this region. Current studies found that most cis-acting sequences sufficient to regulate gene expression resided within 2 kb upstream of the starting codon in *C. elegans* (Levine and Tjian 2003), thus the 2.1 kb promoter sequence of *mks-1* is sufficient to drive transgene expression (Efimenko *et al.* 2005) and for potential transcription factors binding. I compared *pmks-1::gfp* transgene expression patterns in the strain that is generated in this study and the strain that Efimenko *et al.* generated.

Efimenko *et al.* reports diminished *pmks-1::gfp* transgene expression in *daf-19 (m86)* mutant while Williams *et al.* reports abolished *pmks-1::YFP* transgene expression in the mutant (Efimenko *et al*. 2005; Williams *et al*. 2008). Because there is a difference in previous reports, I would like to confirm the observation of *pmks-1::gfp* transgene expression in *daf-19 (m86)* mutant. I generated a strain that carries *pmks-1::gfp* transgene in *daf-19 (m86)* background by crossing the reporter strain with *daf-19 (m86)* mutant strain. I observed transgene expression to understand the effect of *daf-19 (m86)* null-allele on *pmks-1::gfp* transgene expression. Thus, I set up a criterion for genetic screening in my study.

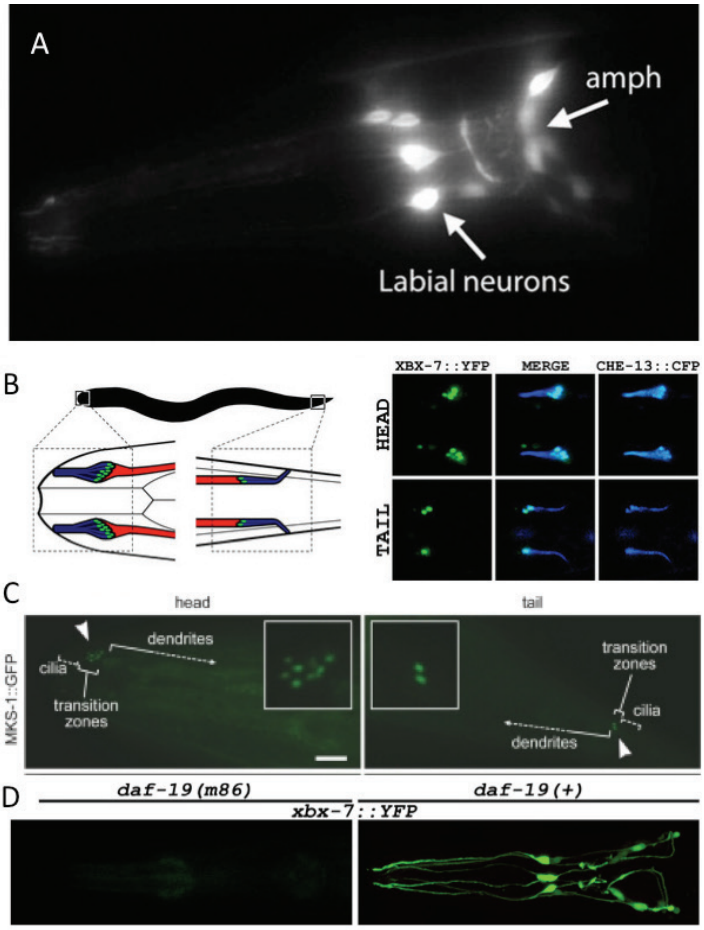


Figure 2.1. Transcriptional expression and translational expression of *mks-1* (a.k.a *xbx-7*) in *C. elegans*

Note. A: *mks-1* promoter driven GFP expression in labial neurons and amphids neurons indicated by arrows (lateral view), phasmids neurons not shown. Images obtained from (Efimenko *et al.* 2005) with permission, doi: 10.1242/dev.01775. B: *mks-1 (xbx-7)* promoter driven YFP expression in transition zone of cilia (dorsal view). Images obtained from (Williams *et al.* 2008). Left: schematic that explains amphid cilia bundles in head and phasmid cilia bundles in tail. Dentrites are labelled in red, transition zone in green, and axoneme in blue. Right: *che-13* promoter driven CFP express along anxoneme and transition zone in cilia, and is used as cilia marker. *mks-1* (*xbx-7)* express at transition zone of cilia in amphid neuron in head (top) and phasmid neuron in tail (bottom). C: GFP-tagged MKS-1 localize specifically in transition zone of cilia in ciliated sensory neurons. Images obtained from (Bialas *et al.* 2009) with permission, doi: 10.1242/jcs.028621. Left: MKS-1 localize to transition zone of cilia at tip of dentrites in head amphid neurons (arrowhead and white box). Right: MKS-1 localize to transition zone of cilia at tip of dentrites in tail phasmid neurons (arrowhead and white box).The transgenic constucts include endogeneous promoter region of *mks-1* and entire coding region of *mks-1*. Scale bar: 5 µm D: *xbx-7::YFP* expression in *daf-19 (m86)* and wildtype background. Images obtained from (Williams *et al.* 2008). Left: *xbx-7::YFP* expression is absent in *daf-19 (m86)* mutant. Right: After outcross the transgnic strain in *daf-19 (m86)* mutant (the left image showed) to *daf-19 (+)* widtype worms, the transgene *xbx-7::YFP* expression is restored.

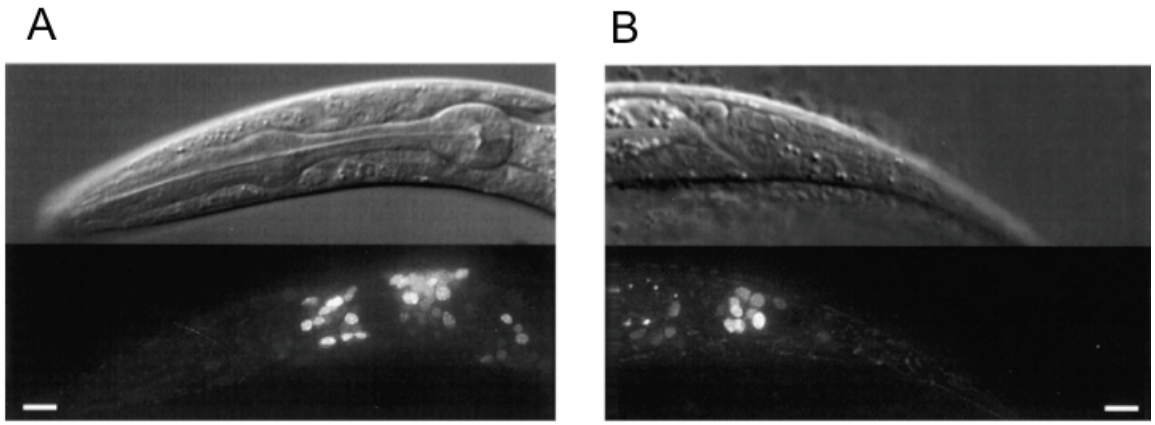


Figure 2.2. *daf-19* promoter driven GFP expression in *C. elegans*.

Note. 2.9 kb *daf-19* promoter and 10 kb *daf-19* genimic sequence were fused to GFP to drive transgenic expression. Expression is seen in all ciliated neurons. Images obtained from (Swoboda *et al.* 2000). Top: Normaski images of *daf-19::gfp* transgenic worm head (A) and tail (B) at L1 stage. Bottom: GFP imgeas of *daf-19::gfp* transgenic worm head (A) and tail (B) at L1 stage. Scale bar: 5 μm

2.2. Materials and Methods

2.2.1. Strains and worm growing conditions

Worms were grown in Easiest Worm Agar plates seeded with OP50. The medium contains 11.8 g Easiest Worm Plate Base Mix (55 g Tris-HCL, 24 g Tris-OH, 310 g Bacto-Peptone, 800 mg Cholesterol and 200 g NaCl) and 36.0 g agar in 2 L ddH_2O (Brenner 1974). The OP50 strain was cultured in Tryptic Soy Broth culture before it was seeded to agar plates. The strain for microinjection is *dpy-5 (e907)* mutant, and was obtained from Baillie lab. The *daf-19* mutant strain is DR86 (*m86*), and was obtained from Swoboda lab. The hermaphrodites N2 strain for backcross is BC9184. The male N2 strain for backcross is BC9183. All worms were growing at 20°C except DR86. DR86 was growing at 15°C. DR86 strain had around 85% dauer formation rate (J. Wang, personal communication). The dauer formation rate of DR86 strain at 15°C was reported as 67% (Swoboda *et al.* 2000).

2.2.2. Generation of transcriptional mks-1 expression construct

2.1 kb intergenic region on the 5' upstream of *mks-1* was amplified from *C. elegans* genomic DNA, because the region was considered as *mks-1* promoter sequences to drive transcriptional expression (Figure 2.3). GFP sequence was amplified from GFP expression plasmids Pd95.75. Jun Wang generated PCR fusion constructs of *mks-1* promoter fused to GFP with standard protocols (Figure 2.4) (Boulin *et al.* 2006).

2.2.3. Microinjection

The linear constructs of *mks-1* promoter fused to GFP sequences were co-injected with *dpy-5* rescue plasmids pcEh361 into *dpy-5 (e907)* mutant worms. Around 20 *dpy-5* mutant worms were injected each time. Zhaozhao Qin kindly injected the worms using Tritech Microinjector Platform (Mello *et al.* 1991; Mello and Fire 1995). I selected F1 progeny with wildtype phenotype. Those worms have successfully injected constructs. I observed GFP expression in F2 generations. I selected wildtype F2

progeny with strong GFP expression in amphids, inner labial neurons and phasmids. Progeny from same injected P0 worms is considered as one line and given same strain name. I got three transgenic worm lines. I named them JNC225 *dotEx225 [pmks-1::gfp;dpy-5(+)]*, JNC226 *dotEx226 [pmks-1::gfp;dpy-5(+)]*, and JNC227 *dotEx227 [pmks-1::gfp;dpy-5(+)]* respectively. To maintain the strains, I picked up wildtype worms, and transferred them to fresh plates, three worms per plates. All the injected lines were frozen in L1-L2 stage in liquid nitrogen. The freezing solution recipe is: 240 ml M9 (contains 3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl and 1 ml of 1 M $MgSO_4$ in 1 L dH_2O), 158 ml S Buffer (contains 5.845 g NaCl and 6.805 g KH_2PO_4 in 1 L dH_2O, PH adjusted to 6.0 with approximately 5-6 ml of 1M NaOH) and 72 ml Glycerol. The ingredients (M9, S Buffer and Glycerol) were autoclaved separately and then mixed by sterile techniques. The sterile freezing solutions were kept at 4^oC.

2.2.4. Integration and Outcross to N2

The JCN225 strain (genotype: *dotEx225 [pmks-1::gfp;dpy-5(+)])* was used for X-ray integration. I prepared a plate of ~200 worms synchronized at young adult stages. I treated these young adults with 145 kV X-ray for 135 seconds. After one-hour recovery, I placed ~40 X-ray treated worms on ~40 large plates and grew them at 23°C. After 24 hours, the X-ray treated worms were killed, leaving around ~50 embryos of F1 generation. Then the F1 were growing at 23°C until the F2 generation grew to L4 stage. With the help from Zhaozhao Qin and undergraduate student John Coo, I picked up ~600 L4-stage worms in F2 generation, one per small plate. After three to four days at 20°C, the F3 generations grew up. At that time, all ~600 plates were screened for integrated lines. In the integrated lines, all F3 progeny have wildtype phenotypes. I also observed plates of all F3 generation worms under a fluorescence dissecting microscope. All F3 worms on the plate had *mks-1* promoter driven GFP expression in head and tail. Worms in these lines have transgenic arrays integrated into chromosome. In total, I obtained three independent integrated lines. These lines were named JNC236 *dotIs236 [pmks-1::gfp;dpy-5(+)]*, JNC237 *dotIs237 [pmks-1::gfp;dpy-5(+)]* and JNC238 *dotIs238 [pmks-1::gfp;dpy-5(+)]*. Because JNC236 has strongest GFP expression, I chose it for following experiments. JNC236 *dotIs236 [pmks-1::gfp;dpy-5(+)]* was backcrossed with N2 strains for four times (Figure 2.5). The backcrossed strain was used for mutagenesis

(see chapter 3 for EMS mutagenesis). The backcrossed strain that carries integrated *mks-1* promoter driven GFP expression reporter was named JNC248 *dotIs248 [pmks-1::gfp;dpy-5(+)]*.

While I used JNC248 reporter strain that stably expresses *pmks-1::gfp* for candidate gene approach experiments (see appendix for details), I discovered that transgene is linked to *zip-5,* a gene on chromosome V. It was hard to isolate double recombinant progeny that carry homozygous *pmks-1::gfp* transgene and homozygous *zip-5 (gk646)* mutant allele, so *pmks-1::gfp* transgene array is inserted on the ChrV in strain JNC248.

2.2.5. Confocal microscopy

I observed the GFP expression of JNC225 and JNC248 strain by confocal microscopy using 40X oil lens. Glass slides with 2% agarose placed on top were prepared for confocal observation. Worms were fixed on the agarose with M9 solution and 5% (300 mM) NaN_2 (1.95 g solid NaN_3 dissolved in 100 ml ddH_2O). M9 solution contains 3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl and 1 ml of 1 M $MgSO_4$ in 1L dH_2O. The worms were observed in the ZEISS confocal microscopy platform.

2.2.6. Cross mks-1 promoter drive GFP reporter into daf-19 mutant background

In order to observe *mks-1* promoter drive GFP expression in *daf-19* mutant background, I introduced *mks-1* promoter driven GFP reporter into *daf-19* mutant by genetic cross. I crossed the *pmks-1::gfp* transgene of JNC248 strain (*dotIs236 [pmks-1::gfp;dpy-5(+)])* with *daf-19 (m86)* mutant of DR86 strain (Figure 2.6). *daf-19 (m86)* allele is the null allele that previous studies also used (Efimenko *et al*. 2005; Williams *et al*. 2008). I obtained a worm line that had heterozygous *mks-1* promoter driven GFP reporter in homozygous *daf-19 (m86)* mutant. Then I observed GFP signal in this worm line, and compared it to that in wildtype transgenic worms (JNC248 strain).

2.2.7. Single Worm lysis and PCR genotyping

To get DNA from a single worm, I ran the single worm lysis program. The lysis buffer was prepared with 10 mM Tris (PH 8.4), 50 mM KCl, 2.5 mM $MgCl_2$, 0.45% Tween 20, and 0.5 mg/mL gelatin. 500 µl lysis buffer was aliquot into each eppendorf tube. 5 µl proteinase K (2 mg/µl) was added to every 500 µl lysis buffer. PCR tubes with 5 µl mix of lysis buffer and proteinase K in each were prepared. A single worm was put into the 5 µl mix in each PCR tube. The tubes with buffer mix and worms were put in -80°C freezer for half an hour so that worm cells were broken in extreme low temperature. The frozen tubes were taken out to room temperature. The tubes were centrifuged for 30 seconds. Then the tubes were put in PCR machine with program: 60°C, 60 min; 95°C, 15 min; keep at 4°C. The 60°C was the working temperature for proteinase K to release the DNA. The 95°C was the temperature to deactivate extra proteinase K. After running the lysis program, the lysis DNA was used as template for PCR genotyping. The lysis DNA was stored in -20°C freezer.

The PCR genotyping aimed to use specific primers to test the existence of mutant and transgene alleles. For the cross of *pmks-1::gfp* transgene to *daf-19 (m86)* mutant (see Figure 2.6), Jun Wang designed primers to detect GFP sequence and *m86* point mutation. The primers to detect GFP sequence are: Forward: 5' GAAACGGCATGACTTTTTCAA 3', Reverse: 5' CTTTTCGTTGGGATCTTTCG 3'. The GFP primers bind to regions inside GFP coding sequences. The GFP primers generated a band of around ~500 bp in GFP transgenic worms, but in wildtype worms the primers do not generate any band. I also used a pair of testing primers to indicate that the PCR worked properly. The test primers' sequences are: Forward : 5' TTATGCAGCGCTCAACATTT 3', Reverse: 5' GGGTCTTGCCACCACTAAAA 3'. The testing primers generated a band of around ~800 bp in all reactions. I used testing primers together with GFP primers in the reaction to genotype GFP. In reactions to genotype GFP, a band of ~500 bp and a band of ~800 bp shows up in worms with GFP transgene, while only one band of ~800 bp shows up in worms without GFP transgene. Because of using testing primers, I confirmed that all PCR reactions were working properly. The primers to detect *m86* point mutation are: Forward: 5' ACCTACCACAGTGCAATACC 3', Reverse: 5' ACCGTTGCGGGGGAAGCGCA 3'. The *m86* forward primer binds to sequences upstream to the point mutation. The first

nucleotide in the *m86* reverse primer is A. Thus the *m86* reverse primer specifically binds to the C -> T mutation and its upstream region. The *m86* primers generated a band of around ~750 bp in *m86* mutant, but the *m86* primers do not generate any band in wildtype worms.

The recipe for PCR reactions with one pair of primers was: Total 10 µl, includes ddH_2O 5.45 µl, 10XTBA 1 µl, $MgCl_2$ 0.4 µl, dNTP 0.25 µl, Forward Primer 0.4 µl, Reverse Primer 0.4 µl, lysis template 2 µl, Taq polymerase 0.1 µl. The recipe for PCR reactions with two pairs of primers was: Total 10 µl, includes ddH_2O 4.65 µl, 10XTBA 1 µl, $MgCl_2$ 0.4 µl, dNTP 0.25 µl, Forward Primer 0.4 µl each, Second Forward Primer 0.4 µl each, lysis template 2 µl, Taq polymerase 0.1 µl. The PCR program for GFP genotyping was: 94°C 10 min; x25 cycles of 94°C 30 s, 58°C 30 s and 72°C 1 min 30 s; 72°C 1 min 30 s; keep at 10°C. The annealing temperature was adjusted to 62°C in *m86* genotyping. PCR products were separated by DNA electrophoresis using 1% agarose gel that contained 0.02% ethidium bromide (10µl ethidium bromide in every 50ml agarose gel). The DNA products were visualized using UV light gel documentation system Geni2. The DNA ladders were Generuler 1 kb or 1 kb plus ladder.

2.2.8. Generation of Males

Males were generated by heat-shock. 60 late L4 worms were placed on 60 petri agar plates, with one worm per plate. Worms were divided into three groups, with 20 worms in each group. The three groups were incubated at 30°C for 5 hours, 5.5 hours and 6 hours, respectively. Then the worms were placed at 20°C incubator. After 3-4 days, the F1 progeny were screened for males. To keep the males, around 8 males were crossed to two L4 hermaphrodites. The hermaphrodites were not treated with heat shock. All the males were maintained in the mating plates, which was seeded with around 1 cm diameter circle of OP50 in the middle.

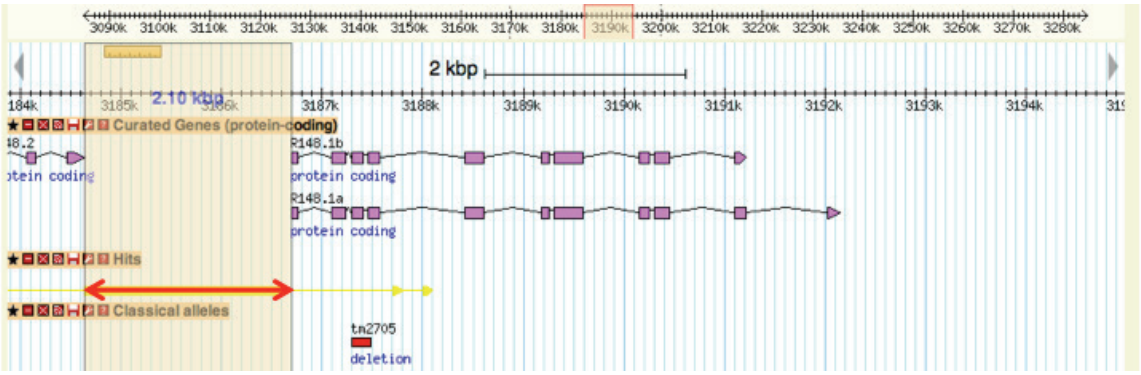


Figure 2.3. Promoter sequences used to generate *mks-1* transcriptional fusion construct

Note. Images were download from Wormbase genome browser (http://www.wormbase.org). The ~2.1kb *mks-1* 5' upstream intergenic sequences is used to generate *mks-1* transcriptional fusion construct (in red bar).

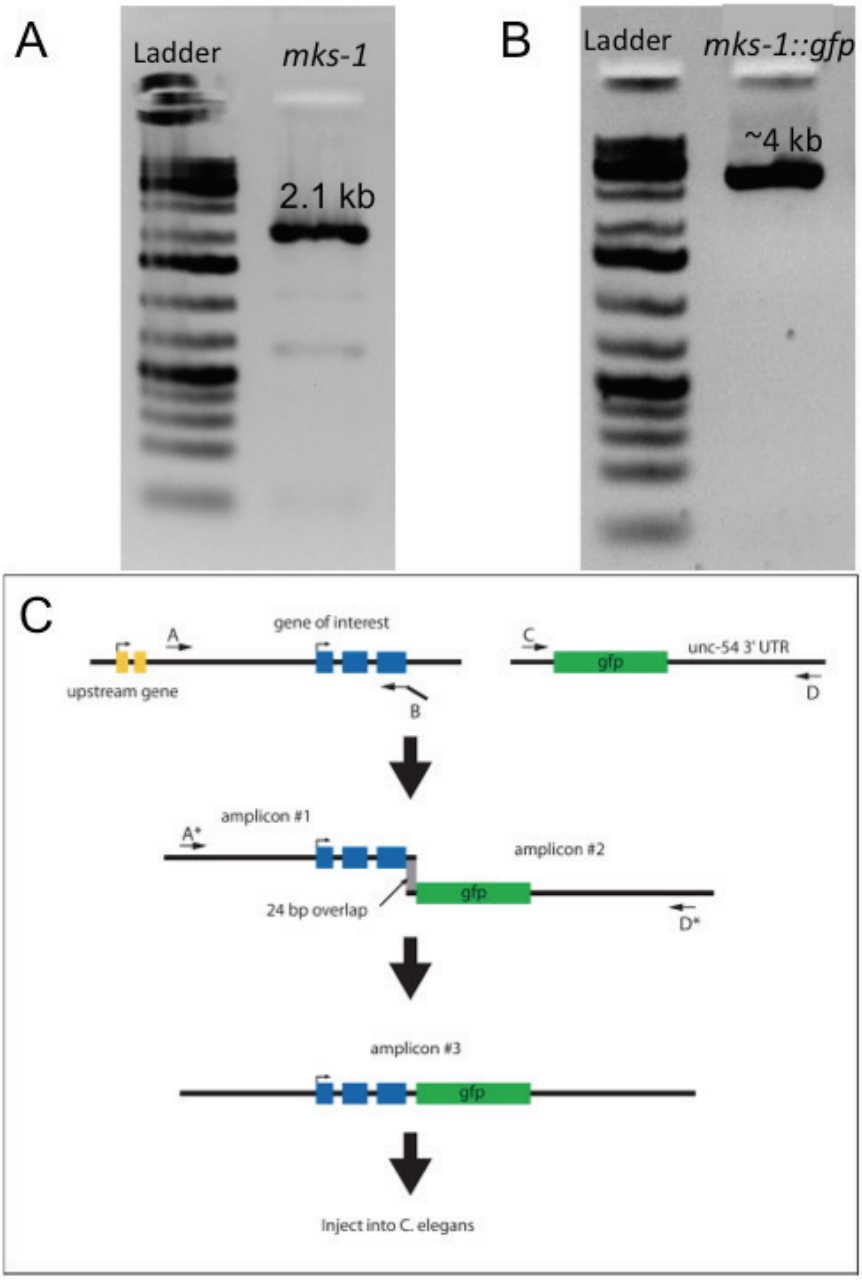


Figure 2.4. PCR fusion of *mks-1* promoter to *GFP* sequence

Note. A: PCR amplicon of ~2.1 kb *mks-1* promoter region. B: Fusion-PCR amplicon of ~2.1 kb *mks-1* promoter fused to ~1.9 kb GFP DNA sequence. The total fusion product length is ~ 4 kb. DNA ladders in A and B are Generuler DNA ladder 1 kb plus. A and B images are obtained from J.Wang. C: Procedure of generating PCR fusion construct. The figure is adopted from (Boulin *et al*. 2006). Primer A and primer B amplify the interested gene regions (amplicon #1). Primer B adds an overlap with GFP coding region in the end. Primer C and primer D amplify the reporter GFP gene coding region and 3' UTR (amplicon #2). Primer A* and primer D* are used to fuse amplicon #1 (interested gene regions) and amplicon #2 (reporter GFP gene). Gray box indicated the overlap regions. The fusion product of interested gene regions fused to GFP (amplicon #3) can be injected to *C. elegans* to observe expression.

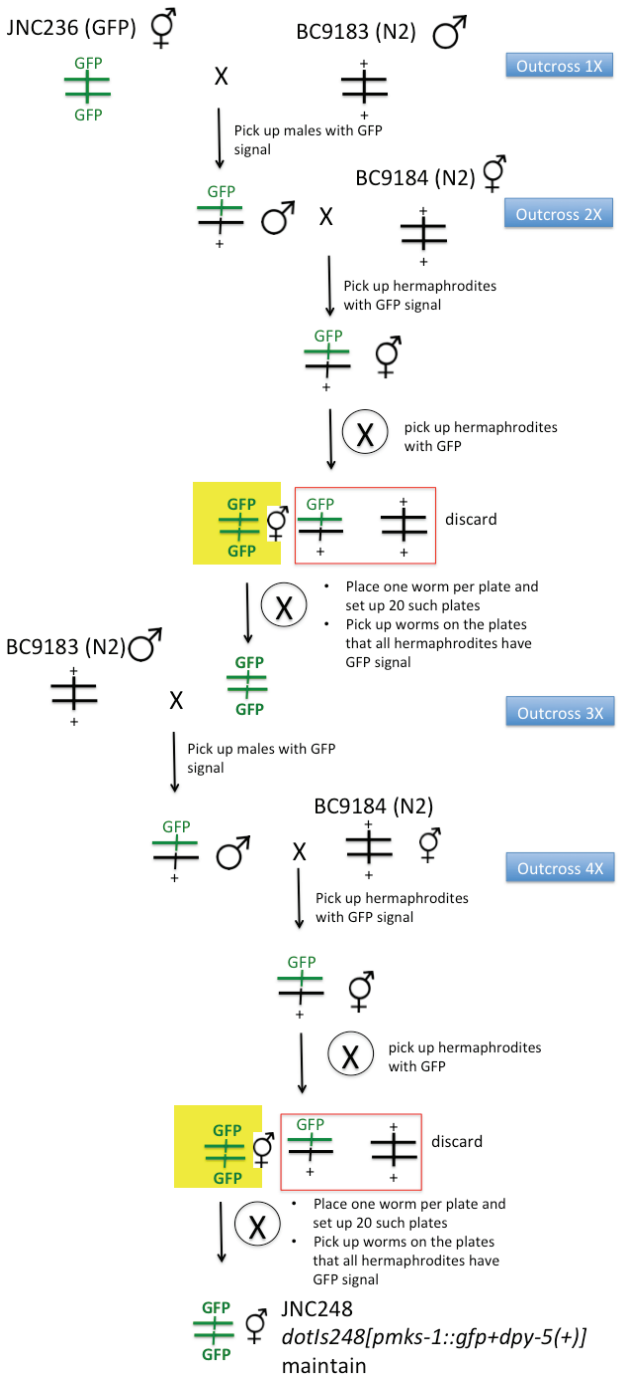


Figure 2.5. Protocol of outcross JNC236 strain to N2 strain

Note. The JNC236 strain is outcrossed to N2 for four times. In outcross 1X and outcross 2X, hetrozygous progeny were picked up. After 2X outcross, homozygous hermaphrodites were isolated from heterozygous prodeny. The 2X outcrossed homozygous hermaphrodites were kept so that if any further cross failed, the outcross did need to be reconducted from 1X. The 2X outcrossed homozygous hermaphrodites were crossed to N2 again for outcross 3X and followed by outcross 4X. After 4X outcross, homozyugous hermaphrodites were isolated from herterozygous progeny. The 4X outcrossed reporter strain's genotype is *dotIs248 [pmks-1::gfp+dpy-5(+)]* and strain name is JNC248.

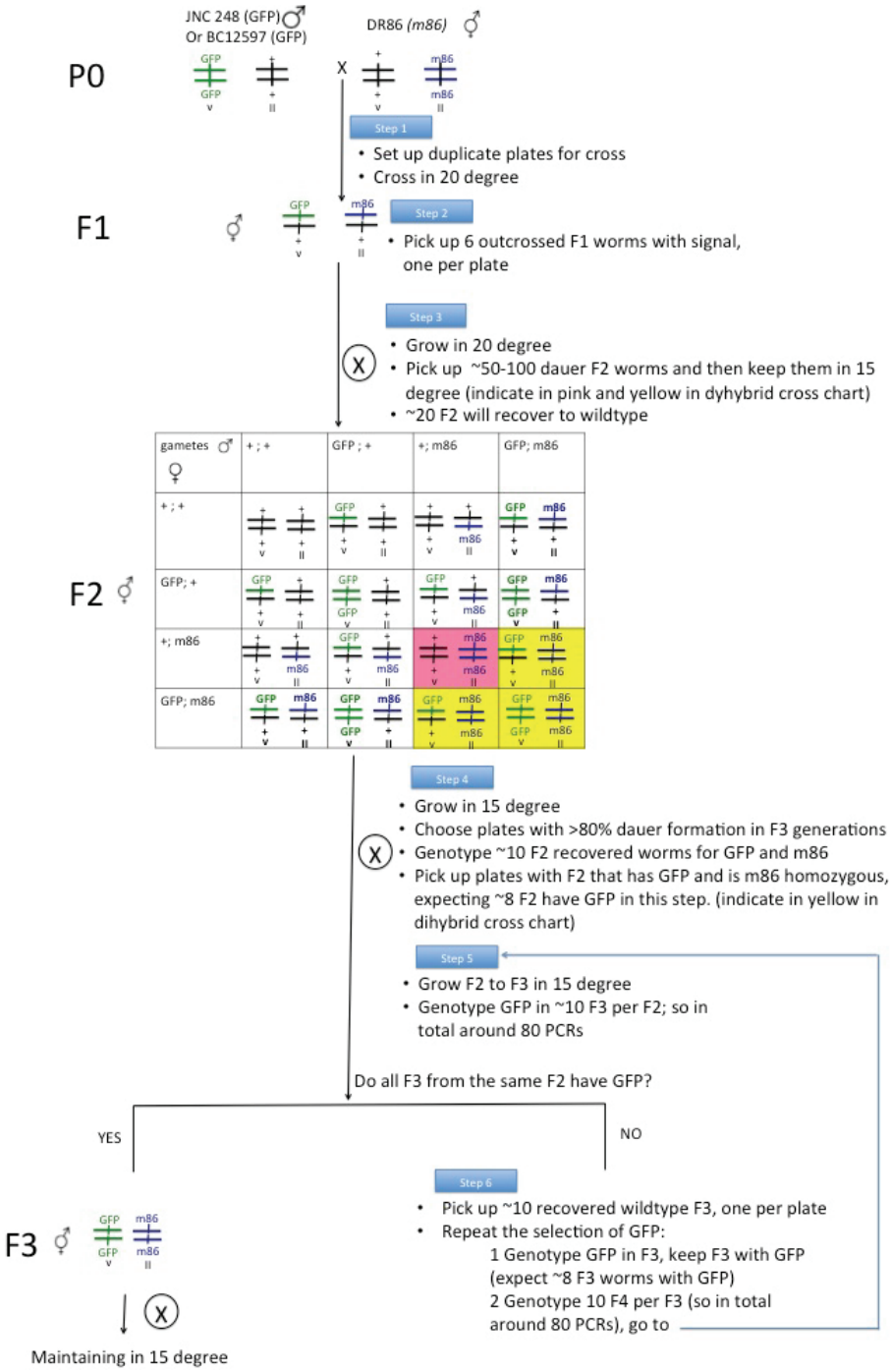


Figure 2.6. Protocol of cross *mks-1* promoter driven GFP reporter into *daf-19 (m86)* mutant

Note. The JNC248 is the GFP reporter strain and DR86 is the *daf-19 (m86)* mutant strain. Heterozygous F1 hermaphrodites were picked up. Homozygous *daf-19 (m86)* mutant were first isolated in F2 generation. Then worms with *pmks-1::gfp* transgenic arrays were selected from the homozygous *daf-19 (m86)* mutant F2 hermaprodites. *pmks-1::gfp* transgene is located on the ChrV because in one genetic cross (see Appendix), the transgene is linked to *zip-5 (gk646)* on the ChrV.

2.3. Results and Discussions

2.3.1. Transgenic mks-1 promoter driven GFP expression in wildtype background

In my study, I generated fusion constructs of 2.1 kb promoter fused to GFP coding sequences. I used the same promoter sequences as previously reported (Efimenko *et al*. 2005; Williams *et al*. 2008). I microinjected *pmks-1::gfp* constructs together with *dpy-5* rescue plasmids into *dpy-5* mutant worms to get transgenic strain that carry *mks-1* promoter driven GFP reporter. In order to get a strain that stably expresses *mks-1* promoter driven GFP, I used X-ray to integrate the extra-chromosomal array of *pmks-1::gfp* constructs and *dpy-5* rescue constructs into chromosome of transgenic worms. Meanwhile, I obtained transgenic strain OE3152 (*ofEx120 [R148.1::gfp;rol-6(su1006)*]) from Swoboda lab, that Efimenko *et al*. generated in previous study. Worms of strain OE3152 have Rol phenotype, and wildtype *pmks-1::gfp* expression. I expected to observe same *pmks-1::gfp* expression patterns in both the strains generated in my study and the strain that Efimenko *et al.* generated.

The expression of GFP in JNC225, JNC248 and OE5132 strain were carefully observed by confocal microscopy using the 40X oil lens. JNC225 strain is the transgenic strain with extra-chromosomal *mks-1* promoter driven GFP reporter. JNC248 strain is the transgenic strain with integrated *mks-1* promoter driven GFP reporter. OE5132 strain is the transgenic strain with extra-chromosomal *mks-1* promoter driven GFP reporter (Efimenko *et al.* 2005).

In transgenic strain JNC225 and JNC248 that were generated in my study, *pmks-1::gfp* expresses in 3 pairs of inner labial neurons, 12 pairs of amphids and 2 pairs of phasmids (Fig 2.7 A-H and Figure 2.8). In transgenic strain OE3152 that were generated by Efimemko *et al.* (Efimenko *et al*. 2005)*, pmks-1::gfp* also expresses in 3 pairs of inner labial neurons, 12 pairs of amphids and 2 pairs of phasmids (Figure 2.7 I-L). There are 6 pairs of inner labial neurons in *C. elegans*. Three pairs are termed IL1 and the other three are termed IL2 (Ward et al. 1975; Ware 1975). IL1 and IL2 cells are close to each other, so it is hard to distinguish whether *pmks-1::gfp* expresses in IL1 or IL2 (Figure 2.8). Further co-localization study, for example, cell staining, is required to identify the

inner labial neurons in *pmks-1::gfp* transgenic worms. The expression patterns are the same in the strains generated in my study and the strain generated in previous study. However, the expression level in phasmids is higher in Efimenko *et al*'s strain OE3152. Although there are variations of expression levels in different transgenic strains, the expression patterns are the identical to previous report. Thus, the transgenic strains JNC248, which stably expresses *pmks-1::gfp* in a wildtype phenotype, is used for EMS mutagenesis and genetic screening.

2.3.2. Transgenic mks-1 promoter driven GFP expression in daf-19 mutant background

In order to confirm the DAF-19 loss-of-function effect on *mks-1* expression, I crossed *pmks-1::gfp* reporter strain (JNC248) to *daf-19 (m86)* mutant strain (DR86).I picked up heterozygous *m86* and *pmks-1::gfp* F1 progeny. I isolated homozygous *m86* progeny that carry *pmks-1::gfp* reporter from self-crossed progeny of heterozygous F1 generation. This is a *pmks-1::gfp* heterozygous and *m86* homozygous clone. I confirmed that this clone is GFP heterozygous and m86 homozygous by genotyping progeny of this clone. I randomly picked up 10 self-fertilized progeny from this clone. All of progeny had *m86* allele but only 8 of them had GFP construct (Figure 2.9).

Then I observed the *mks-1* expression in *daf-19 (m86)* mutant in progeny of this clone that carry heterozygous *pmks-1::gfp* transgenic arrays. I expected to observe absent or diminished *pmks-1::gfp* expression in *daf-19 (m86)* mutant.

I observed that neuronal GFP signal is absent in *daf-19 (m86)* mutant (Figure 2.10). I increased confocal exposure time for observation, but still I could not observe expression in labial neurons, amphids, and phasmids in *daf-19 (m86)* mutant (Figure 2.10 K and L), where *pmks-1::gfp* expresses in those cells in wildtype (Figure 2.10 G and H). When I increased exposure time, non-neuronal GFP signal is shown in images' background in both wildtype and *daf-19 (m86)* mutant due to GFP emission under long exposure time (Waters 2009). Under long exposure time, I did not observe non-neuronal GFP signal in worms with *pmks-1::gfp* transgene that Efimenko *et al* generated (OE3152 strain). Different constructs were used to generate different transgene array in this study and Efimenko *et al.* study. Efimenko *et al.* used GFP expression plasmids while I used

PCR fusion constructs in this study. However, it is yet to known whether using PCR-fusion constructs as transgene resulted in detectable GFP signal in all worm body under high exposure time. Still, I concluded that *pmks-1::gfp* neuronal expression is absent in *daf-19 (m86)* mutant, which is same as reported by Williams *et al.* (Williams *et al*. 2008).

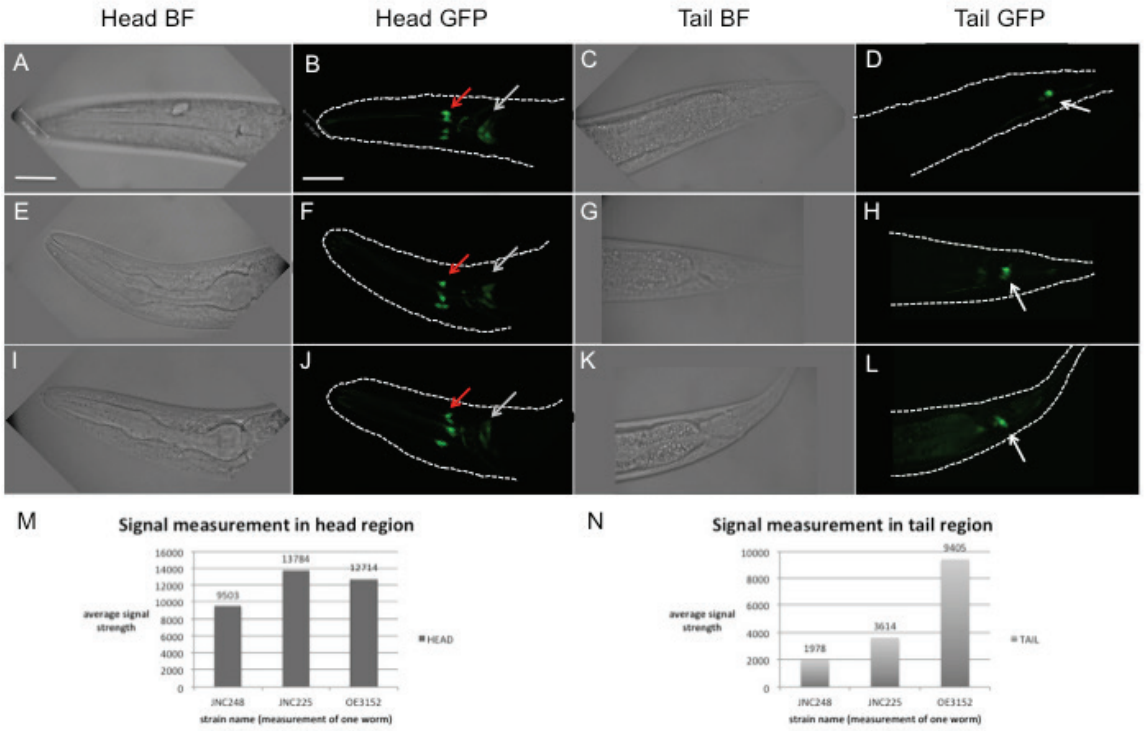


Figure 2.7. ***mks-1*** **promoter driven GFP expression in strain OE3152 (transgenic strain), JNC225 (transgenic strain) and JNC248 (integrated transgenic strain)**

Note. A-D: extra-chrmosomal *mks-1* promoter driven gfp reporters express in labial neurons, amphids and phasmids in OE3152 *ofEx120[R148.1::gfp;rol-6(su1006)*]. E-H: extra-chromosomal *mks-1* promoter driven gfp reporters express in labial neurons, amphids and phasmids in JNC225 *dotEx225[pmks-1::gfp;dpy-5(+)]*. I-L: intra-chromosomal *mks-1* promoter driven gfp reporters express in labial neurons, amphids and phasmids in JNC248 *dotIs248[pmks-1::gfp;dpy-5(+)]*. Inner labial neurons are indicated by red arrowheads, amphids by grey arrowheads and phasmids by white arrowheads. In order to best display the GFP signal, different strains were observed under different settings (i.e. exposure time, sensitivity and contrast enhancement). All GFP images are confocal stacks of whole worm body from lateral view. Scale bar: 25 μm. To compare the GFP signal levels, worms from different strains (one worm per strain) were also observe under same settings and measured their signal strength by counting the signal pixels. For worms measuremed in M and N, GFP exposure was at 182 ms. M: measurement of head region signals in different strains under same confocal settings. The head expression level is slightly higher in transgenic strain JNC225 and OE3152 that carry extra-chromosomal reporters than strain JNC248 that carry intra-chromosomal reporters. N: measurement of tail region signals in different strains under same confocal settings. The tail expression level is slightly higher in strain JNC225 than in strain JNC248. The tail expression level is much higher in strain OE3152. Expression level differences can only be detected by confocal microscopy.

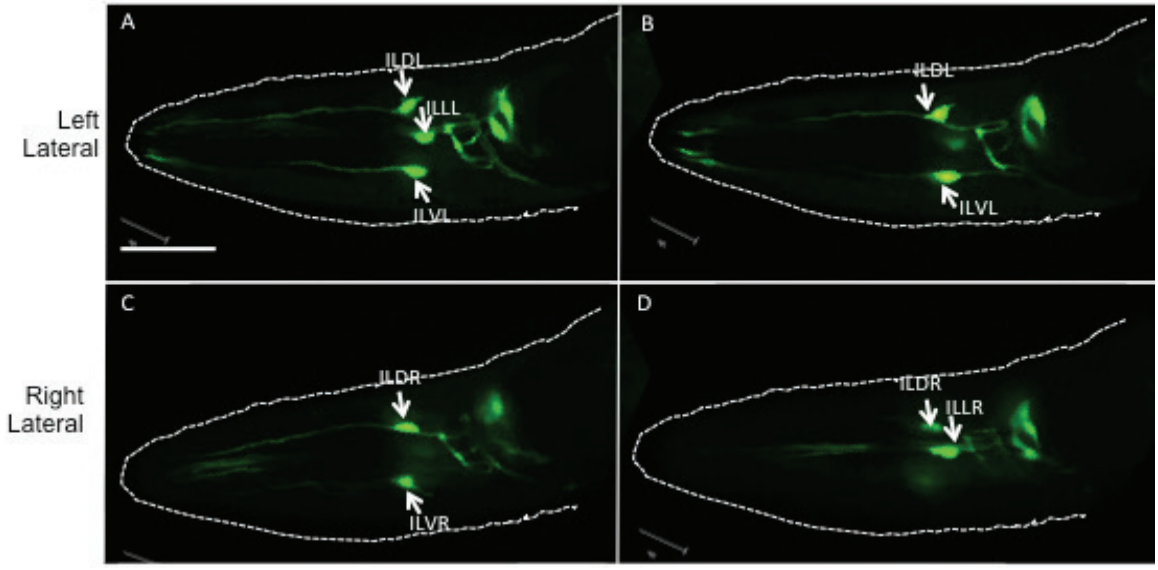


Figure 2.8 Expression of *pmks-1::gfp* in inner labial neurons in wildtype (transgenic strain JNC248)

Note I scanned each layers of the projected confocal image of *pmks-1::gfp* wildtype transgenic worm's head region. 6 neuron cell bodies are identified, although it is yet to be known whether these neurons are IL1 or IL2. A,B : In left lateral side, three cell bodies, of neuron ILDL, ILLL, ILVL (unknown whether IL1 or IL2), are indicated by white arrowheads. C, D: In right lateral side, three cell bodies, of neuron ILDR, ILLR, ILVR (unknown whether IL1 or IL2), are indicated by white arrowheads. Exposure is taken at 182 ms. Brightness is set at x6.5 so the whole cell bodies are clearly displayed. Scale bar: 25 μm.

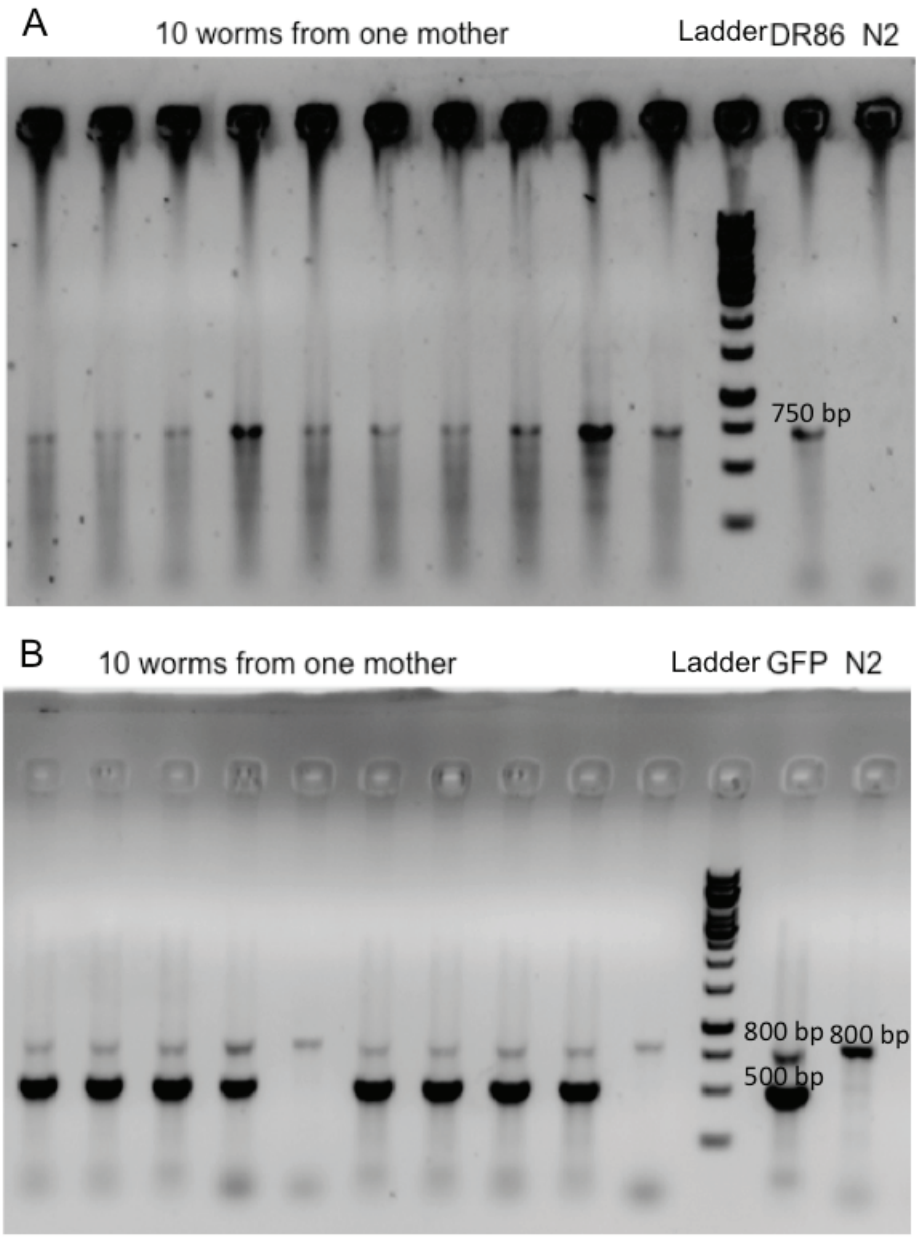


Figure 2.9. DNA electrophoresis result showing the *m86* homozygous and *pmks-1::gfp* heterozygous clone

Note. The figure shows how to determine that a worm clone is homozygous *m86* and heterozygous *pmks-1::gfp*. 10 progeny of a worm clone were lysated to collect DNA in ten tubes. Then 10 lysates were tested with two rounds of PCR (A and B). A: The primers generate one ~ 750 bp in *m86* mutant and none band in wt. Genotyping of 10 progeneis showed 10 *m86* specific band ~750 bp, so ten progeny are from a *m86* homozygous mother. B: The primers generate two bands, ~500 bp and ~800 bp in worms with *pmks-1::gfp* constructs. The primers generate one ~800 bp band in wt. Genotyping of 10 progeneis only showed 8 GFP specific band ~500 bp, so ten progeny are from a *pmks-1::gfp* heterozygous mother. This homozygous *m86* and heterozygous *pmks-1::gfp* reporters line was used to observe *mks-1* promoter driven GFP signal in *m86* mutant. DNA ladder: Gene ruler 1 kb ladder.

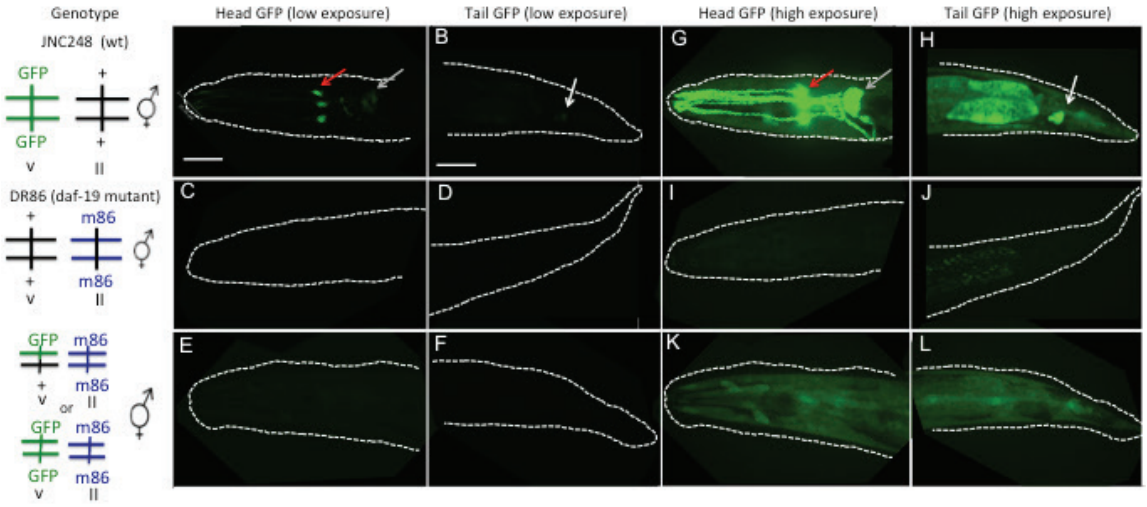


Figure 2.10. ***mks-1*** **promoter driven GFP expression in** ***daf-19 (m86)*** **mutant**

Note. Images A-F were taken under low exposure time, 394 ms A,B: In wildtype strain that carries *pmks-1::gfp* reporters, GFP signal is detected in inner labial neurons (indicated by red arrowhead), amphids (indicated by grey arrowhead) and phasmids (indicated by white arrowhead). C, D: In *m86* mutant strain that does not carry reporters, no signal is detected. E, F: In *m86* mutant that carries *pmks-1::gfp* reporters, GFP signal is absent. A, C, E were taken under same sensitivity. B, D, F were taken under same sensitivity that is a bit higher than A, C, E, because reporter gene expression is higher in head region than tail region (see Fig 2.8 M, N). Images G-L were taken under high exposure time, 768 ms. G, H: In wildtype strain that carries *pmks-1::gfp* reporters, GFP signal is detected in inner labial neurons (indicated by red arrowhead), amphids (indicated by grey arrowhead) and phasmids (indicated by white arrowhead). Because exposure time is high, the signals are saturated. I, J: In *m86* mutant strain that does not carry reporters, no signal is detected. K, L: In *m86* mutant that carries *pmks-1::gfp* reporters, *mks-1* expression in labial neurons, amphids and phasmids is suppressed because no GFP signal is detected in these regions. Non-neuronal expression is introduced by *pmks-1::gfp* reporter genes, which can also be seen in the background of image G and H. All GFP images are confocal stacks of whole worm body from lateral view. Scale bar: 25 µm.

2.4. Conclusion

First, I prepared reporter strain (JNC248 *dotIs248 [pmks-1::gfp;dpy-5(+)]*) that would be used for genetic screening (see Chapter 3) and candidate gene approach (See Appendix). I confirmed that *pmks-1::gfp* expresses in 3 pairs of inner labial neurons, 12 pairs of amphids and 2 pairs of phasmids. The transgene expression resembles *in-vivo mks-1* gene expression. My observation of wildtype *mks-1* promoter driven GFP expression pattern is consistent with previous report (Efimenko *et al.* 2005; Williams *et al.* 2008). It is possible that different dosage of injected constructs caused the difference of transgene expression level. In my study, the concentration of injected *pmks-1::gfp* fusion constructs is 15 µg/ µl. And the concentration of transgene constructs that Efimenko *et al*. used is not revealed (Efimenko *et al*. 2005).

Second, I confirmed *mks-1* expression level in *daf-19 (m86)* mutant, although previous reports are conflicting. Efimenko *et al.* first observed diminished *pmks-1::gfp* expression in *daf-19 (m86)* mutant in 2005 (Efimenko *et al*. 2005). However, in 2008 Williams *et al*. described *pmks-1::gfp* expression in *daf-19 (m86)* mutant as absent (Figure 2.1 D) (Williams *et al*. 2008). Both of them injected *pmks-1::gfp* constructs into *daf-19 (m86)* mutants. Efimenko *et al.* used clones of GFP plasmids that are inserted with *mks-1* promoter sequences. Slightly differently, Williams *et al.* used PCR fusion constructs of *mks-1* promoter sequences fused to YFP coding sequences. In my study, I also used *pmks-1::gfp* transcriptional fusion constructs. I injected the constructs with *dpy-5* rescue vectors into *dpy-5 (e907)* mutants to generate a wildtype transgenic strain. I generated a transgene-integrated strain from injection strain. Then I crossed this transgene-integrated strain with *daf-19 (m86*) mutant to get a strain that carries both alleles. I observed absent *pmks-1::gfp* neuronal expression in *daf-19 (m86)* mutant. My observation is consistent with report from Williams *et al.* (Williams *et al*. 2008). In Efimekno *et al.* study and my study, different injection constructs were used. It could be possible that cloning plasmids have higher expression than PCR fusion constructs level and thus it is hard to shut down transgene expression in transgenic worms that carry expression plasmids.

I come to a criterion for genetic screening based on my observations of *pmks-1::gfp* expression in wildtype and *daf-19 (m86)* mutant. I would look for mutant with reduced GFP signal since deleting DAF-19 resulted in loss of *pmks-1::gfp* neuronal expression.

Chapter 3.

Identifying *mks-1* transcription factors through genetic screening

3.1. Background

Effimenko *et al.* found X-box sequence, which is for RFX characteristic DBD binding, in *mks-1* upstream intragenic regions (Efimenko *et al*. 2005). The *mks-1* expression is DAF-19 dependent based on the evidence that *mks-1* promoter driven GFP expression is absent in *daf-19* null-mutant in (Williams *et al*. 2008) and current study. Several lines of evidence reveal that *mks-1* is regulated by other transcription factors together with DAF-19. First, DAF-19^{C} isoform is expressed in all ciliated neurons, while *mks-1* only express in subset of ciliated neurons. Efimenko *et al*. predicted existence of co-factor to DAF-19 in regulating ciliary genes that only express in subset of ciliated neurons (Efimenko *et al*. 2005). This group of ciliary genes has function in specific cells in cilia formation (Thomas *et al*. 2010; Choksi *et al*. 2014). Second, evidences of co-regulation in cilia exist. For example, RFX3 is a co-factor to FOXJ1 to augment FOXJ1 dependent transcription in regulating motile cilia formation in mouse airway cells (Didon *et al*. 2013). Third, it is reported that some cell-specific transcription factors work upstream or downstream of RFX to regulated ciliary gene expression. In *C. elegans*, transcription factor FKH-2 functions downstream to DAF-19 to regulate ciliary gene *kap-1* expression in AWB neurons (Mukhopadhyay *et al*. 2007b). Thus, I hypothesized that *mks-1* is regulated by a group of transcriptional factors including DAF-19. In order to find transcription factors in addition to DAF-19, I applied different approaches, including genetic screening and a candidate gene approach. In the candidate gene approach, I introduced *mks-1* promoter driven GFP reporters into different mutants. The candidate genes encode already known transcription factors in *C.*

elegans (See Appendix). However, I spent most time and energy on genetic screens to find mutants of new candidate transcription factor(s).

In my study, I applied EMS mutagenesis and genetic screens based on Brenner's methods (Brenner 1974). EMS (Ethyl methanesulfonate, CH3SO3C2H5) is an alkylating mutagen that adds an ethyl group to guanine to form O^6-ethylguanine which pairs with thymine, thus causing GC to AT substitutions (Brookes and Lawley 1961; Coulondre and Miller 1977). It generates mutations in a rate of 1 in 2000 genes to 1 in 20 genes, depending on the concentration and treatment time (Gengyo-Ando and Mitani 2000). My study used GFP as reporters for genetic screening. Because *C. elegans* body is transparent, it is ideal to monitor gene expression in live worms by using fluorescence transgene reporters. It is a mature technique to use GFP as reporters in genetic screens to search for changes in cell positioning or gene expression levels. For example, Troemel *et al*. generated worms carrying *str-2* promoter fused to GFP transgene array. *str-2* encodes an odorant receptor that expresses in one of the bilaterally symmetric AWC neurons. They successfully identified mutants that express *str-2::GFP* in both AWC neurons or no AWC neurons. Their screens help identifying regulators in the AWC neuron-specific signaling pathway (Troemel *et al*. 1999). Doitsidou *et al.* applied similar genetic screen strategy to search for mutants defective in dopaminergic cell differentiation. They use GFP reporter fused with promoter of *dat-1* to label dopaminergic neurons. *dat-1* encodes a plasma membrane dopamine transporter which exclusively express in dopaminergic neurons. They applied two ways for screens, either manually or by using COPAS Biosort system (worm sorter) to select mutants of subtle alterations in GFP expression. Their study revealed high efficiency in using worm sorter for GFP-labeled genetic screens (Doitsidou *et al*. 2008).

In this thesis study, I looked for mutants with reduced *pmks-1::gfp* expression. I expected to find mutants that specifically affect *mks-1* or ciliary gene expression. Because DAF-19 is a known transcription factor of *mks-1*, I expected to isolate *daf-19* mutant that has absent or reduced *pmks-1::gfp* expression. Because EMS generates random mutations in a rate of 1 in 1000 genes (Brenner 1974), in this study I expected to obtain EMS-induced mutations in some genes from large gene families such as *unc* or *dpy*. I expected to see mutants of Unc (uncoordinated) phenotype because *unc* genes belong to one of the largest gene family in *C. elegans*. I expected that EMS mutagen

would also hit genes in a wide spectrum of pathways that affecting cell development, cell cycles, apoptosis, cell positioning and so on. Thus, I expected to see mutants that are larva-arrest, lethal, sterile, vulvaless or malformed in different organs. If these mutations locate in genes that upstream to *mks-1* regulating pathway, I would expect to observe reduced or abolished GFP expression in these larva-arrest, lethal, sterile, or malformed mutants.

3.2. Materials and Methods

3.2.1. Strains

I used the reporter strain JNC248 for first round of EMS mutagenesis and genetic screen (generation of JNC248 described in Chapter 1, Figure 2.5 and Figure 2.7). In second round of EMS mutagenesis, I used a double-reporter strain JNC285 with *pmks-1::gfp* and *ppgp-12::gfp* reporters. JNC285 strain was generated by genetic cross. The transgenic *ppgp-12::gfp* strain was obtained from Baillie lab. The *pgp-12* promoter driven GFP reporter was integrated into chromosome. Then the *ppgp-12::gfp* integrated strain was backcrossed to N2 strain for four times to clean background mutations. The cleaned *ppgp-12::gfp* integrated strain is named JNC296 *dotIs296 [ppgp-12::gfp;dpy-5(+)]*. I crossed *dotIs248 [pmks-1::gfp;dpy-5(+)]* transgenic worms (strain JNC248) with *dotIs296[ppgp-12::gfp;dpy-5(+)]* transgenic worms (strain JNC296) (Figure 3.1). I selected the outcrossed heterozygous F1 generation hermaphrodites. In F2 progeny, 1 out of 16 worms have two homozygous GFP reporters. The F2 generation worm with both homozygous *pmks-1::gfp* and *ppgp-12::gfp* reporters was selected. This strain was named JNC285 (genotype: *dotIs248 [pmks-1::gfp;dpy-5(+)]; dotIs296 [ppgp-12::gfp;dpy-5(+)])*. I observed the expression of GFP in JNC285 under a confocal microscope. Worms in JNC285 strain have *pmks-1::gfp* expression in inner labial neurons, amphids and phasmids; and *ppgp-12::gfp* expression in the excretory cell (Figure 3.2). The N2 male is BC9183. All experiments were conducted at 20°C, unless otherwise noted for temperature sensitive strains.

3.2.2. EMS mutagenesis and genetic screening

In order to find mutants of candidate transcriptional factors of *mks-1*, I applied two rounds of genetic screens (Figure 3.3). In first round, worms carrying *pmks-1::gfp* reporter (JNC248 strain) were mutagenized. Around 200 young adult worms were treated with 50 mM EMS for 4 hours and then recovered in room temperature for an hour. The EMS treated worms were put on large plates. They grew until all the worms on plates were starved at L1-L2 stage. I used non-clonal F2 screening. When screening F2 generation, I was able to get homozygous mutants. Moreover, by F2 non-clonal screen, I do not need to pick up individual F1 clones, which saves time and energy. If F1 progeny is heterozygous, 1/4 of the F2 progeny is homozygous. The F2 and F3 generation starved worms were transferred to fresh plates by chunking agar and putting it on new plates. After the F2 and F3 generations grew up to L4 stage on fresh plates, I screened them under a dissecting fluorescent microscope. Worms with altered GFP expression (i.e. reduced or abolished expression) in amphids, phasmids and inner labial neurons were picked up.

The procedures of second round of genetic screen were very similar to the first rounds. There were three major improvements (Figure 3.4). First, the mutagenized worms have two reporter transgene *ppmks-1::gfp* and *ppgp-12::gfp* (JNC285 strain). *pmks-1:::gfp* represents the expression of target gene *mks-1*. *ppgp-12::gfp* represents a transmembrane coding gene expression in the excretory cell. There are different regulating pathways between *pgp-12* and ciliary gene *mks-1*. *pgpg-12::gfp* is a control reporter because mutants affecting *ppgp-12::gfp* expression are in regulators unrelated to ciliary-gene specific regulation. Second, I used different ways to transfer F2 or F3 generations. In first round of genetic screening, I transferred almost all starving F2 or F3 generations to fresh plates. Because each F1 generation mutant is estimated to lay ~150 embryos, most F2 generation mutants on the same plate are segregates of the same F1 generation mutant. Thus I do not need to screen all F2 progeny. In second round, only starving F2 or F3 generations on 9 pieces of agar in 1 cm^2 size of each original plate were transferred to fresh plates. Third, the phenotypes in the second round of screening were altered GFP expression in amphids, phasmids or inner labial neurons, but unchanged expression in excretory cells. The changed screen criterion enables me to pick up mutants that specifically affect ciliary gene expression in neurons. Because

cell-specific transcription factors could be discovered from the screen (Mukhopadhyay *et al.* 2007b), the altered GFP includes reductions in overall signal strength as well as reductions in specific neurons (i.e. missing expression in certain neurons). After screening, I selected only one mutant from each P0 plate because multiple mutants on same plates are likely to be segregates from same mutant F1 progeny, considered the number of discovered mutants and number of progeny that a F1 generation mutant could lay. The selected mutants were characterized as recessive, dominant or transgenic-array defective, with genetic crossings.

3.2.3. Maintenance and selections of EMS mutants

All mutants were maintained at 20°C unless any mutants were identified as temperature sensitive. ~3-5 worms with mutant GFP phenotypes were picked up under a fluorescence dissecting microscope for maintaining. After maintaining three generations, I selected mutant lines that have stably altered *pmks-1::gfp* expression. These mutant lines have 100% penetrance. Mutants with 100% penetrance were kept for regular maintaining by picking up ~3-5 worms each time under a normal dissecting microscope. The worms were kept for more than ten generations. If GFP signal in some mutant strains recovered to wildtype level after ten generations, they were not studied. Mutant strains with 100% penetrance were characterized.

3.2.4. Characterizations of EMS mutants

In order to know whether the mutants carried dominant mutations, recessive mutations or transgenic-array defective, I crossed the mutant hermaphrodites with N2 male (Figure 3.5). I observed GFP signals in all F1 generation worms under a fluorescence dissecting microscope. The F1 outcrossed males were carefully observed under a confocal microscope. The mutants and wildtype *pmks-1::gfp* transgenic worms were used as controls for confocal observation. If F1 generation hermaphrodites and males both have wildtype GFP signal, the mutant is autosomal recessive. If F1 generation hermaphrodites have wildtype GFP signal, and F1 generation males have altered GFP signal (mutant phenotype), the mutant is X-linked recessive. If the F1 generation worms all have altered GFP signal (mutant phenotype), the mutant hermaphrodites was crossed with wildtype *pmks-1::gfp* transgenic male worms to

determine whether the mutation is dominant or transgenic-array defective (Figure 3.6). Again, I observed GFP signals in all F1 generation worms under a fluorescence dissecting microscope. And GFP signal in the F1 generation males were carefully observed under a confocal microscope. If F1 generation worms all have altered GFP signal (mutant phenotypes), the mutant is dominant. If F1 generation hermaphrodites have wildtype GFP signal, the mutant has defective transgenic arrays. Depending on where the defective arrays are located, the F1 generation males would have wildtype GFP signal if the arrays are located on the autosome, while the GFP signal is altered (mutant phenotype) in F1 generation males if the arrays are located on the X-chromosome.

3.2.5. Preliminary Genetic analysis of EMS mutants

After two rounds of genetic screens, I isolated a X-linked recessive mutation (*dot14*) that suppresses *mks-1* expression in subset of labial neurons. In order to preliminarily understand whether the mutations affect transcription factors of *mks-1* or affect neuron cell differentiation, I labeled labial neurons in the mutant strain by genetic cross (Figure 3.7A). I used the strain BC12597 (*dpy-2(e907) I; sIs12174 [rCes C47A10.6::GFP+pCeh361])*. The strain has stable *srab-12* promoter driven GFP expression in labial neurons, phasmids, and scattered nonchemosensory neurons along worm body (Hunt-Newbury *et al.* 2007). SRAB-12 is a transmembrane serpentine receptor that belongs to class AB family (Chen *et al*. 2005). *srab-12* expresses in inner labial neurons although it is not a ciliary gene. Thus *srab-12* and *mks-1* are regulated by different pathways. Transgenic worms that stably expresses *psrab-12::gfp* (BC12597) is obtained from Baillie lab. I used heat shock to generate male worms carrying *psrab-12::gfp* transgene. The male worms carrying *srab-12* promoter driven GFP reporter was crossed with mutants that carried *mks-1* promoter driven GFP reporter and mutant alleles. The F1 generation males were carefully observed under a confocal microscope.

In order to know whether the mutations affect other ciliary gene expression, I introduced *tub-1* reporter transgene into mutants. Gene *tub-1* encodes TUBBY homolog that is important for fat storage and life span (Mukhopadhyay *et al.* 2005). Gene *tub-1* expressed in ciliated neurons including amphids and phasmids in *C. elegans* (Efimenko *et al.* 2005). In Efimenko *et al.* prediction, *tub-1* and *mks-1* belong to the same group of

ciliated genes that function specifically in subset of ciliated neurons. It is highly possible that gene in this group is regulated by some tissue specific transcription factors. I would like to know whether the mutation affect regulation on this group of genes.

In order to track *tub-1* expression, I generated a *tub-1* promoter driven tdTomato transgenic strain. Jun Wang kindly perfumed the PCR fusion of *ptub-1::tdTomato* constructs. I microinjected the constructs of PCR-fused *tub-1* promoter to tdTomato sequence and *dpy-5* rescue plasmids pcEh361 into *dpy-5 (e907)* mutant worms. I generated the strain of stable *ptub-1::tdTomato* expression using X-ray introduced integration (See Chapter 2 methods). The transgenic strain of integrated *tub-1* promoter driven tdTomoato was named JNC247 (*dotIs247 [ptub-1::gfp+dpy-5(+)]*). Then I used heat shock method to generate male transgenic worms of *tub-1* promoter driven tdTomato expression (See Chapter 2 methods). I crossed the *ptub-1::tdTomato* transgenic worms with the *dot14* mutants, and screened for F1 generation males (Figure 3.7B). F1 males were carefully observed under confocal microscope.

In order to know whether the mutants have Dyf phenotype (Dye-filling defect), I applied dye-filling assay to mutant worms. The DiI (1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate) could enter worm phasmids and 6 pairs of amphids neurons through the exposed ciliated endings (Herman 1984). The DiI is observed as red under red fluorescence. Abnormal Dye-filling (Dyf) phenotypes, i.e. failure to take dye in these neurons, could be indication of abnormal cilia, although some ciliary mutants still possessed non-Dyf phenotype (Perens and Shaham 2005; Schafer *et al*. 2006). To perfume dye-filling assay, I washed one plate of worms of L4-to-young adult stage with 1ml M9. The M9 with worms were added to 1.5 ml eppendorf tubes. The tubes were centrifuged at 1200 rpm for 1 min. The supernatant was removed. The worm pellets were kept in the tube. To remove bacteria, the worm pellets were washed with 1 ml M9. Then the tubes were centrifuged at 1200 rpm for 1 min again, and the supernatant was removed again. The wash was repeated twice. Clean worm pellets were collected in tubes. 1 ml M9 and 5 µl DiI (2 mg/ml) were mixed with worm pellets in tube. The tube was wrapped with tinfoil. Worms in tube were incubated at room temperature on rocker for three hours. Then the tubes were centrifuged at 1200 rpm for 1 min, and the supernatant was removed. To remove any left DiI, the worm pellets were washed with 1ml M9. Then the tubes were centrifuged at 1200 rpm for 1 min. Most

supernatant were removed, while only 100-200 µl M9 were left in tube, mixed with worm pellets. Worms in M9 were transferred to plates seeded with OP50. The plates were covered with tinfoil. The worms were left on the plates for an hour to recover. Then I checked all worms under a fluorescence dissecting microscope to observe whether the DiI went into worms phadmids and amphids. I randomly picked up 10 worms, and placed them on 2% agarose pad on glass slide with 10 µl M9. To fixed worms to its position, I added final concentration 10 mM of NaN_3 to the agarose pad (Shaham 2006). The worms were observed under a confocal microscope to check what neurons took DiI. The wildtype reporter strain JNC285 (*dotIs248 [pmks-1::gfp;dpy-5(+)] + dotIs296 [ppgp-12::gfp;dpy-5(+)])* was used as a control.

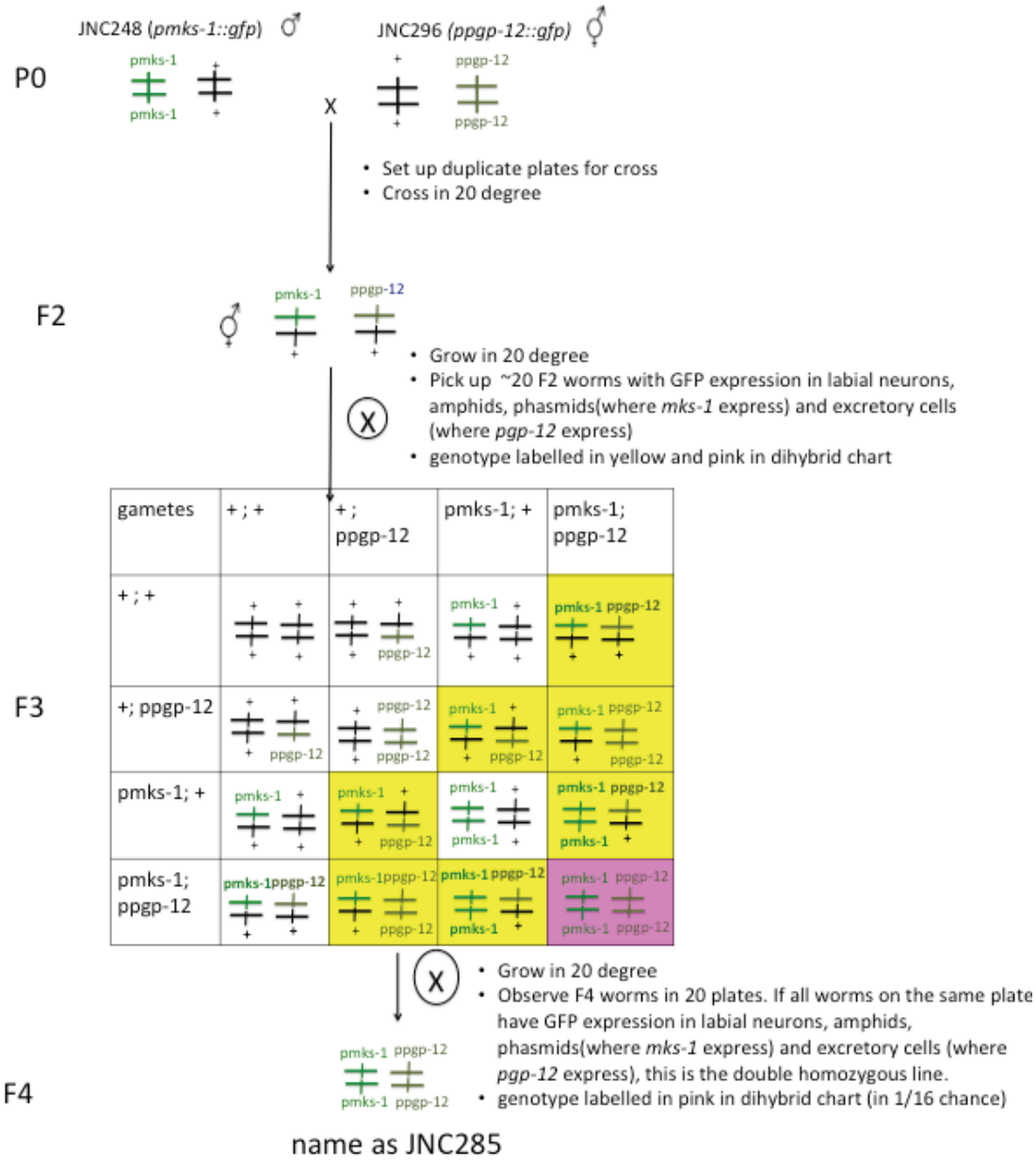


Figure 3.1. Genetic crossing process to generate transgenic worms with double GFP reporters of *pmks-1::gfp* and *ppgp-12::gfp*

Note. The flow chart shows the process to generate JNC285 strain by genetic crossing. Genotypes of F3 generation are shown in the dihybrid chart. In F3 generation, worms with expression in labial neurons, amphids, phasmids and excretory cells were picked up, and placed in separate plates, one worm per plate. F4 generations were observed. Plates with all worms having expression in labial neurons, amphids, phasmids and excretory cells were selected as lines with homozygous *ppgp-12::gfp* and *pmks-1::gfp* reporters' alleles.

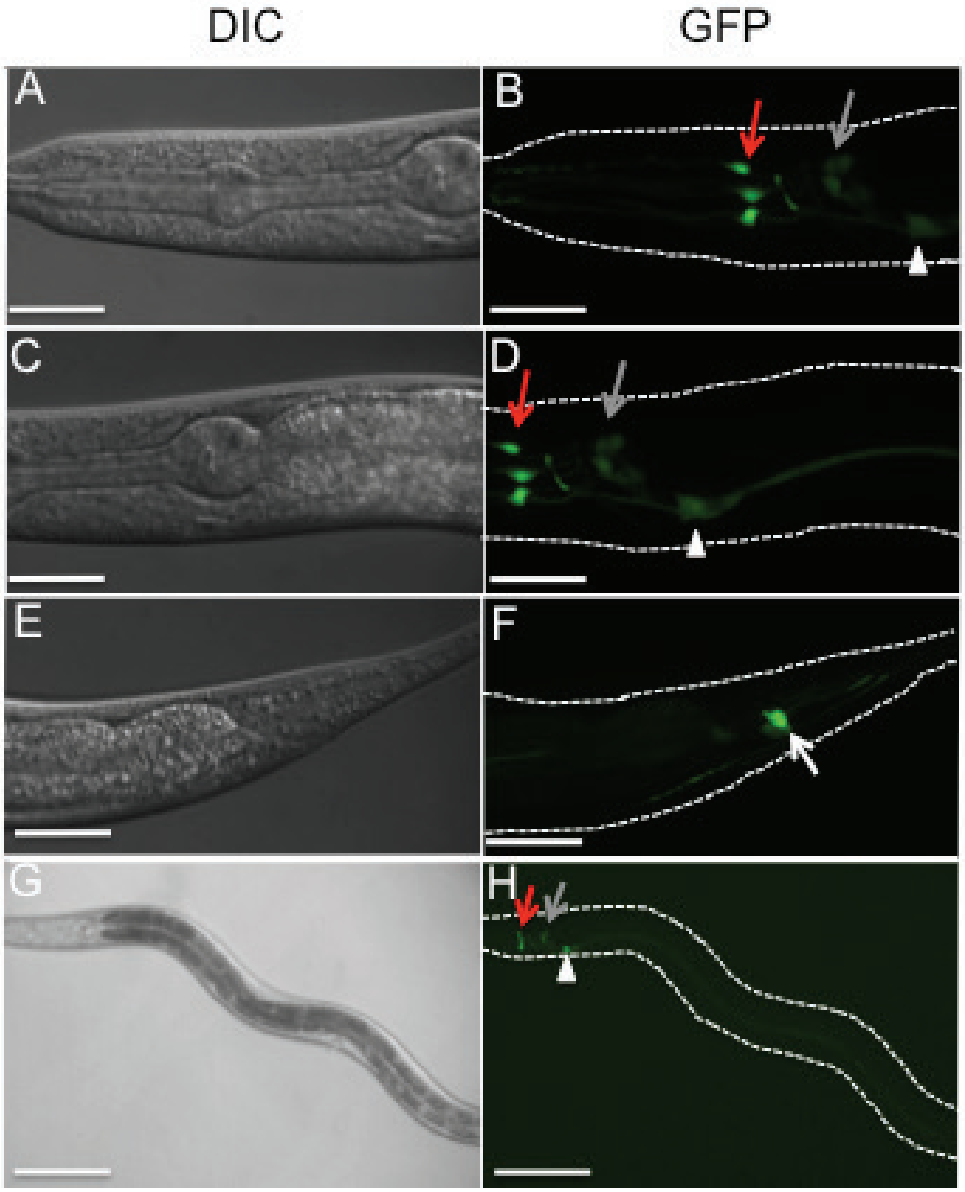


Figure 3.2. GFP expression pattern in transgenic worms with double GFP reporters of *pmks-1::gfp* and *ppgp-12::gfp* (strain JNC285)

Note. Images show confocal observation of transgenic worms of JNC285 strain (*dotIs248[pmks-1::gfp;dpy-5(+)]; dotIs296[ppgp-12::gfp;dpy-5(+)]*). A, B, C,D: *mks-1* promoter driven GFP expresses in labial neurons (red arrowheads) and amphids (grey arrowheads). *pgp-12* promoter driven GFP expresses in excretory cell (white triangles). Scale bar: 25 µm. E,F: *mks-1* promoter driven GFP expresses in phasmids (white arrowhead). Scale bar: 25 µm. G, H: *mks-1* promoter driven GFP expresses in labial neurons (red arrowheads) and amphids (grey arrowheads). *pgp-12* promoter driven GFP expresses in excretory cell (white triangles). Scale bar: 100 µm. All images were taken at exposure 200 ms. To best display the expression patterns, auto-contrast enhancement were applied to all images.

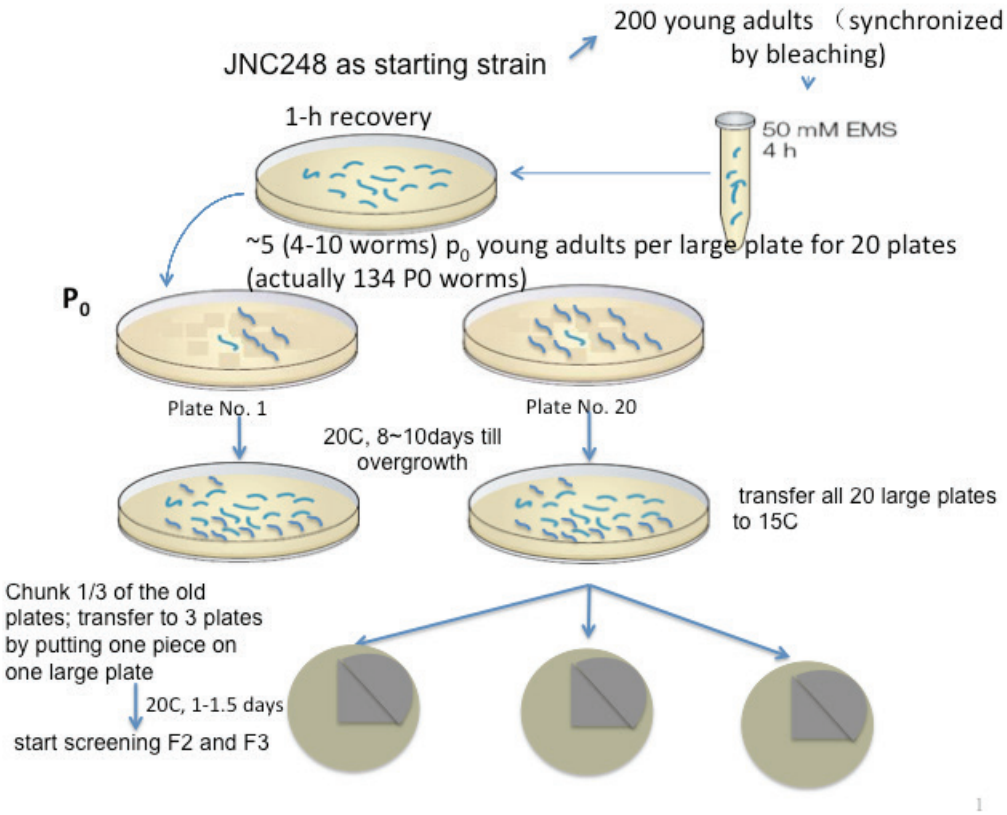


Figure 3.3. Protocol for first round of EMS mutagenesis and genetic screening.

Note. The protocol is modified based on a protocol in Jorgensen and Mango's review (Jorgensen and Mango 2002). Diagram was drawn by Jun Wang and modified by Ting Zhang.

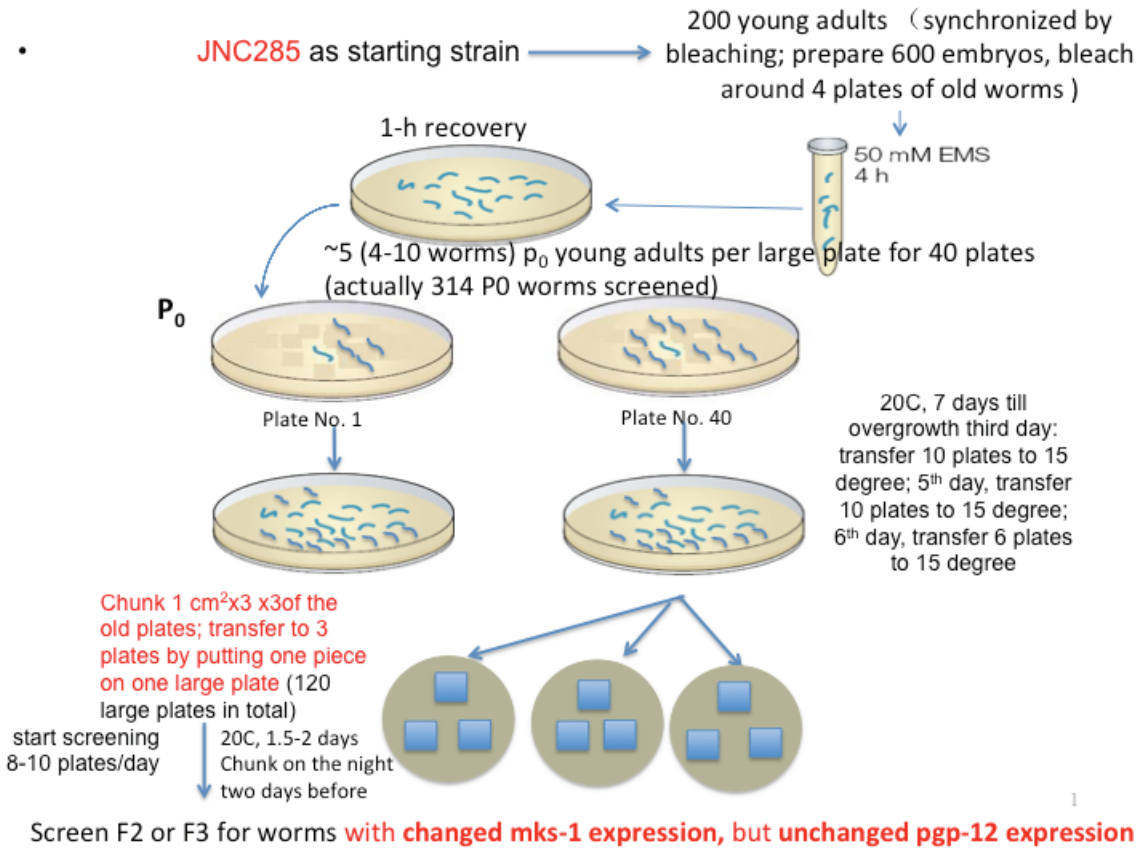


Figure 3.4. Protocol for second round of EMS mutagenesis and genetic screening.

Note. The protocol is modified based on first round of genetic screening. Three major changes are highlighted as red in the protocol. Diagram was drawn by Jun Wang and modified by Ting Zhang.

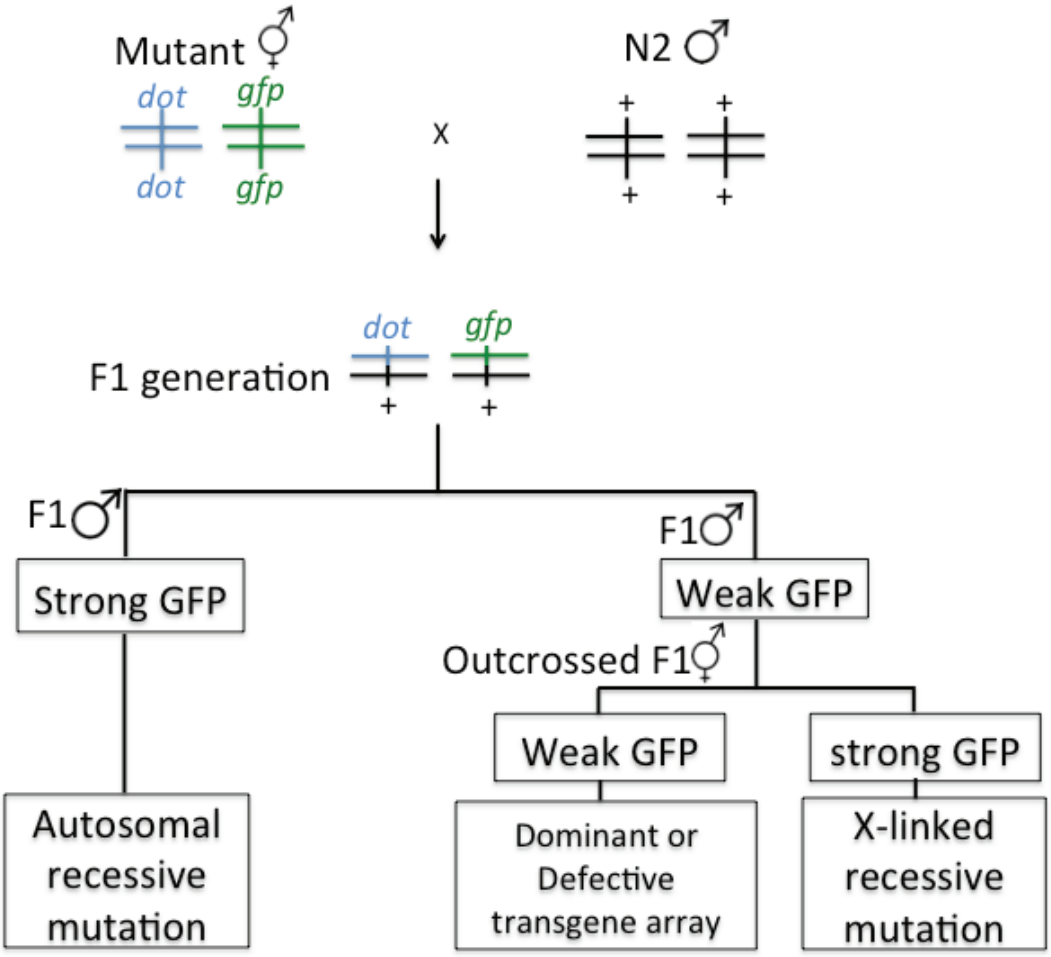


Figure 3.5. Cross mutants to N2 males to characterize mutant allele's traits

Note. I crossed mutant hermaphrodites to N2 males. I observed GFP expression in outcrossed F1 generation. The P0 worms for cross are transferred to new plates after first 24 hours. Progeny on new plates are observed. Because self-crossed embryos are laid first, most self-crossed progeny are not observed in this way. Outcrossed F1 generation's genotype is heterozygous in mutant allele and transgene allele. All males are from outcross. If all male F1 progeny have strong expression, the mutation is autosomal recessive. If all male F1 progeny have weak expression, then F1 hermaphrodites are observed. Most self-crossed progeny are excluded from plates after transferring, and also they have weak GFP. If most F1 hermaphrodites have strong GFP, these are outcrossed F1 hermaphrodites, and the mutation is X-linked recessive. If all F1 hermaphrodites have weak GFP, I crossed mutants to wildtype males that carry *pmks-1::gfp* reporter to determine whether the mutants have dominant mutations or defective transgenic arrays (see Figure 3.6).

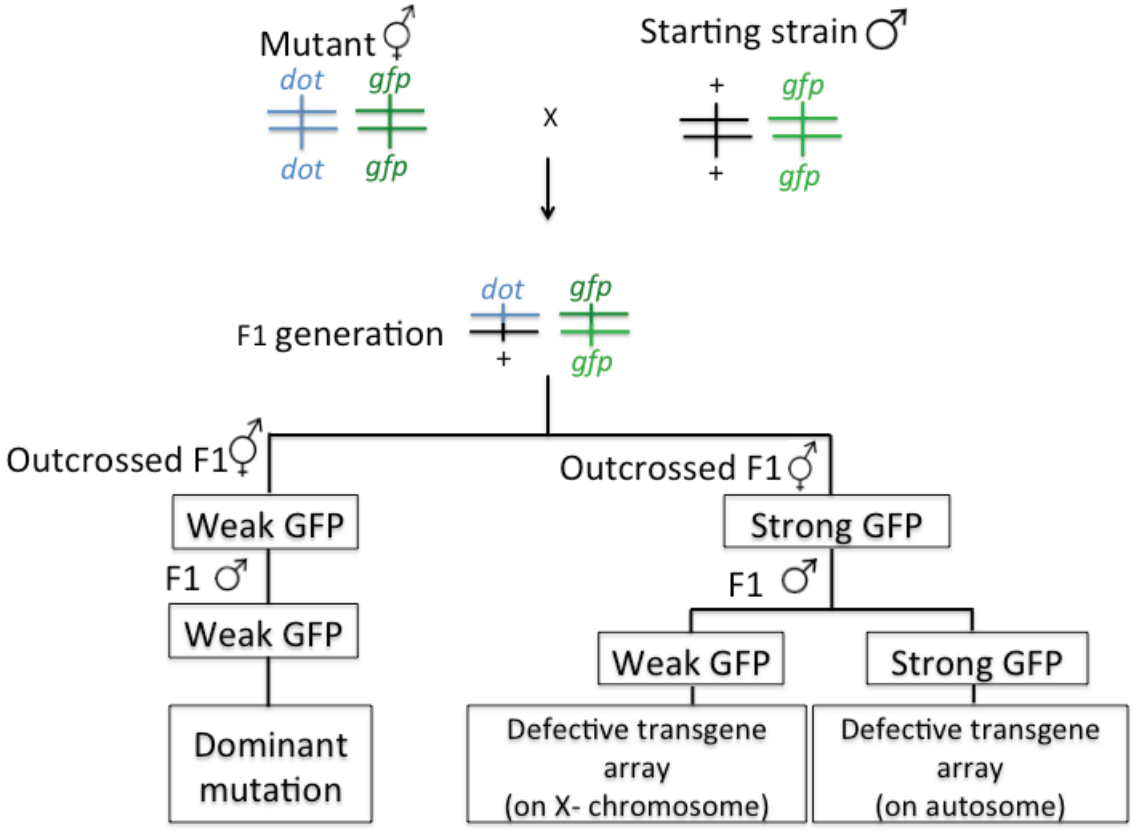


Figure 3.6. Cross mutants to wildtype *pmks-1::gfp* transgenic worm (strain JNC248) to determine whether the mutants have dominant mutation or defective constructs.

Note. I crossed mutant hermaphrodites to *pmks-1::gfp* wildtype transgenic males. I observed GFP expression in outcrossed F1 generation. The worms for cross are transferred to new plates after first 24 hours. Progeny on new plates are observed. Because self-crossed embryos are laid first, most self-crossed progeny are not observed in this way. Outcrossed F1 generation's genotype is heterozygous in mutant allele but homozygous in transgene allele. One copy of transgene is intact, obtained from starting strain. One copy of transgene is from mutants. Most self-crossed progeny are excluded from plates after transferring, and also they have weak GFP. If all F1 hermaphrodites have weak GFP, mutants are dominant. If most F1 hermaphrodites have strong GFP, these are out-crossed F1 hermaphrodites. The mutants have defective transgenic arrays (integrated array of transgenic reporter genes) in this case, because one intact copy of transgenic arrays complement the defective transgenic arrays in mutants. The defective transgene array in mutants could be caused by two possibilities: the transgene size shrinks due to non-homologous recombination after generations, or chromatin enzymes silenced the transgene expression by histone modification. Depending on the male F1 progeny's phenotype, we could know whether the transgene is located on the X chromosome or not. If male F1 progeny also have strong GFP, transgene is located on the autosome. If male F1 progeny have weak GFP, transgene is located on the ChrX.

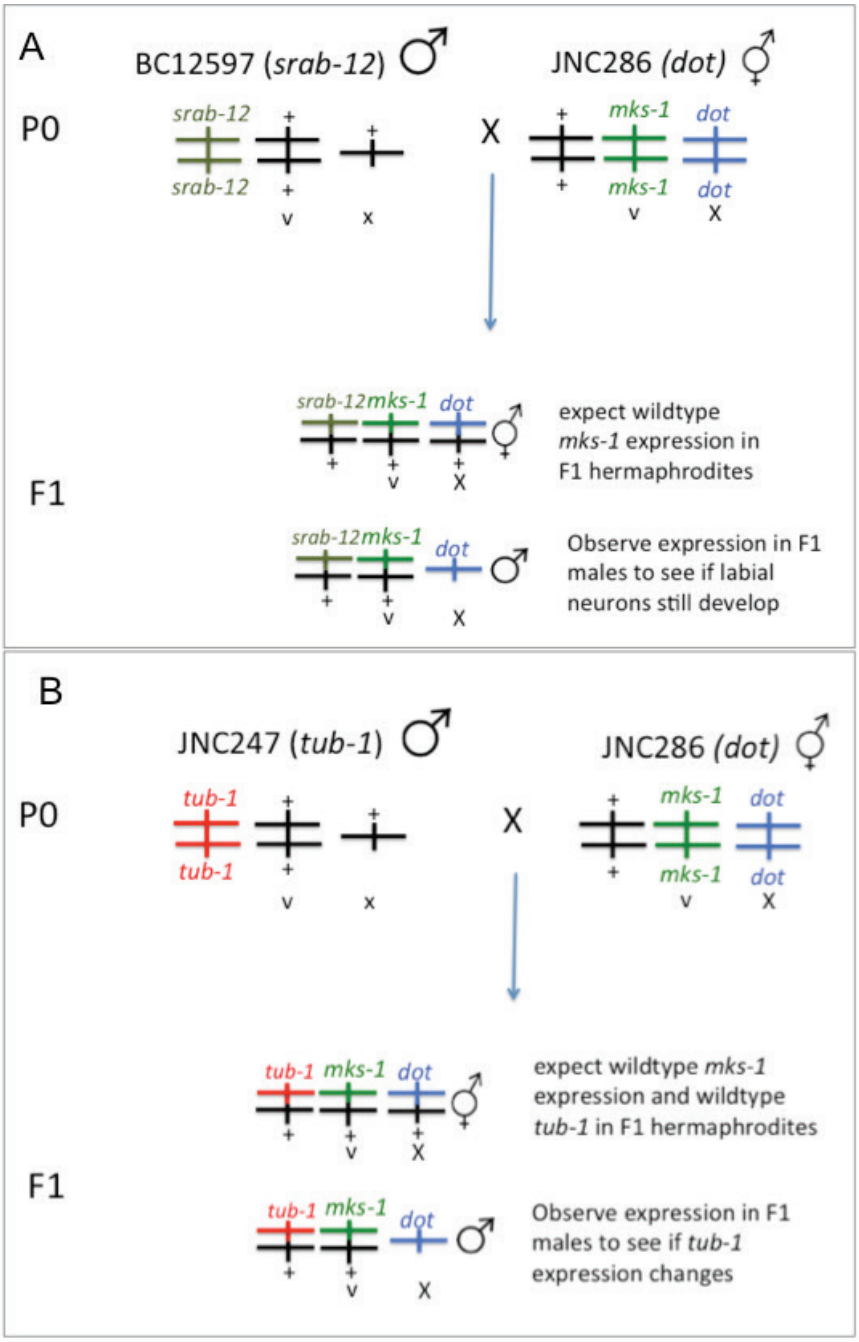


Figure 3.7. Protocol for preliminary analysis of *dot14* mutant

Note. In preliminary analysis of JNC286 mutant, the mutant hermaphrodites were crossed to males carrying *srab-12* promoter driven GFP reporter (A) or males carrying *tub-1* promoter driven tdTomato reporter (B). Because *dot14* is located on the X chromosome, the F1 generation males had mutant phenotype of *mks-1* expression. To determine if the mutations affect labial neuron differentiation, and if the mutations affect ciliary gene *tub-1* expression, I observed phenotypes of labial neuron expression and *ptub-1::tdTomato* expression in F1 generation males.

3.3. Results and Discussions

3.3.1. Two characterized mutants (JNC260,JNC262) from first round of genetic screen have defective tansgene array

I used the strain JNC248 with stably *pmks-1::gfp* expression as a reporter strain for screens. Transgenic worms in JNC248 strain (*dotIs248 [pmks-1::gfp;dpy-5(+)]*) was treated with 50 mM EMS for mutagenesis. I manually screened under a fluorescence dissecting microscope for F2 or F3 generation worms with altered GFP expression in head and tail regions. I aimed to pick up mutants of transcription factors that regulate *mks-1* or ciliary gene specifically. I estimated the screening genome size based on the number of EMS treated P0 worms. Based on the fact that each P0 worms have estimated 150 progeny and that *C. elegans* genome is diploid, I used the calculation that genome size = the number of P0 worms x150 x2 for estimation. The F1 genome size is calculated because in non-clonal F2 screens, F2 genome is inherited from F1 genome, which has the mutations. The first round of genetic screen covered the genome 40,200 times.

After screening, I immediately froze all viable mutants at -80 °C for long-term storage. I am not able to maintain and freeze mutants that have lethal progeny or larva-arrest. Meanwhile, homozygous viable mutants were maintained for three generations. Mutants with low-penetrance and heterozygous mutants were not further studied because it is hard to characterize these mutants as recessive or dominant. With so many mutants on hand, I focus on homozygous viable mutants that are easy to characterize. 12 homozygous viable mutants were selected and carefully observed for GFP expression by confocal microscopy (Table 3.1). The mutants have reduced GFP expression level in either head, tail neurons, or both (Figure 3.8). Although *pmks-1::gfp* expression is reduced in mutants, the *pmks-1::gfp* transgene expresses in same neurons in both wildtype and mutant worms, including inner labial neurons, amphids and phasmids (Figure 3.9). Then I chose one mutant from the same P0 plates, because mutants from the same P0 plates are likely to be segregates from the same F1 mutant progeny. I chose 8 mutants. The eight mutants were maintained for more than ten generations (i.e. three months). After three months, GFP expression level recovered to wildtype for some mutants (Table 3.1). It is yet to be known what caused the GFP signal

to recover to wildtype in these mutants. One possibility is that the transgenic arrays' expression in these mutants was suppressed by epigenetics modification, but the modification disappeared after a few generations. I did not study mutants that recovered wildtype GFP expression (Figure 3.10). Mutant lines 19-2-1 and 3-3-1 have stable mutant phenotypes. These two mutant lines were from different P0 worms. Mutant line 3-3-1 is named strain JNC262 (genotype: *dot7; dotIs248 [pmks-1::gfp+dpy-5(+)]*). Mutant line 19-2-1 is named strain JNC260 (genotype: *dot6; dotIs248 [pmks-1::gfp+dpy-5(+)])*

Table 3.1. Details of homozygous viable mutants selected from first round of genetic screening

Category	P0 plate numbers	Amount of P0 on plate	Mutant serial number	Mutant phenotype under dissecting microscope	Phenotype after 10 generations	Characterization
Details	3	6	3-3-1 (JNC262)	No signal	Weak almost no signal	Defective transgenic arrays
	19	8	19-2-1 (JNC260)	No signal	Weak almost no signal	Defective transgenic arrays
			19-3-1	Weak	Relative weak	Not studied
	32	7	32-1-1	Weak	Relative weak	
	23	5	23-2-1	Weak	Relative weak	
	18	7	18-1-2	Weak only in phasmids	Weak only in phasmids; lost	
	27	9	27-3-5	Weak	Relative strong	
			27-3-2	Weak	Relative strong	
	16	5	16-3-2	Weak	Relative strong	
	4	7	4-2-2	Weak	Relative strong	
			4-2-3	Weak only in phasmids	Relative strong	
	22	4	22-3-2	Weak	Relative strong	

Note. The 3-3-1 and 19-2-1 mutants were used for following study. 3-3-1 was named JNC262 strain. 19-2-1 is named JNC260 strain.

To characterize whether the mutations are recessive, dominant or the mutants have defective transgenic arrays, JNC260 (*dot6; dotIs248 [pmks-1::gfp+dpy-5(+)]*) and JNC262 (*dot7; dotIs248 [pmks-1::gfp+dpy-5(+)]*) mutant hermaphrodites were crossed to

N2 males. I observed outcrossed F1 generations under a fluorescence dissecting microscope. All F1 generation worms have reduced GFP expression phenotype. The F1 generation males were carefully observed under confocal microscope. GFP signal in F1 male progeny is reduced (Figure 3.11 and Figure 3.12). Thus, either the mutations are dominant or the transgene expression is defective. The *dot6* mutant (strain JNC260) and *dot7* mutant (strain JNC262) hermaphrodites were then crossed to wildtype male *pmks-1::gfp* transgenic worms. By cross with wildtype males that carry *pmks-1::gfp* transgene, intact copies of *pmks-1::gfp* transgene was introduced to F1 progeny. I observed that all F1 generation males and outcrossed F1 generation hermaphrodites had wildtype GFP expression. Confocal observation of F1 generation males was also conducted. GFP signal in F1 male progeny is as strong as wildtype male transgenic worms (Figure 3.11 and Figure 3.12). Thus, I concluded that *dot6* and *dot7* mutants have defective transgenic arrays. It might be caused by non-homologous recombination of transgenic arrays during meiosis, which result in loss of transgene copy numbers (J. Wang, personal communication). It might also be caused by chromatin enzyme-induced histone modification, which reduces transgene expression at the epigenetics level (Dr. N. Hawkins, personal communication).

3.3.2. One characterized mutant (JNC286) from second round of genetic screen carries X-linked recessive mutations

In my first round of genetic screen, I isolated two mutants but they had defective transgenic arrays. Doitsidou *et al.* also pick up mutants that carry mutations in transgenic arrays, which silence the expression of transgenic arrays (Doitsidou *et al*. 2008). However, I used transgenic worms with multiple copies of transgenic arrays (approximately 20 copies) in this study. Considering the EMS mutation rate and the length of *C. elegans* genome, it is very hard to mutate most transgene copies to generate a reduced expression phenotype. In the two mutants from first round of screen in this study, the transgene array is likely to be silenced by modification in epigenetics level. Thus, I introduced a second reporter, *pgp-12* promoter fused to GFP coding sequences, in second round of screen. *pgp-12* encodes a transmembrane protein that belongs to P-glycoprotein subclass of ATP-binding cassette (ABC) transporter family (Zhao *et al*. 2005). The gene *pgp-12* expresses in the excretory cell (Zhao *et al*. 2005;

Zhao *et al*. 2007). I introduced *ppgp-12::gfp* reporter to genetic screens based on three reasons. First, *pgp-12* is not regulated by any ciliary gene-specific transcription factors. Mutants with unchanged *ppgp-12::gfp* expression but altered *pmks-1::gfp* expression are highly likely to carry mutations in ciliary-gene specific transcription factors. Second, *pgp-12* and *mks-1* are under a common regulation network that regulates all gene expressions and cell development. Using *ppgp-12::gfp* as a second reporter, I would avoid picking up mutants that affect common regulation pathways, which is not the interest of this study. Third, epigenetics modification on transgenes that silences *pmks-1::gf*p expression could have effect on *ppgp-12::gfp* at the same time. I could avoid picking up mutants with epigenetics modifications of the transgenes.

I generated transgenic worms that carry independent insertions of *pmks-1::gfp* and *ppgp-12::gfp* reporters (strain JNC285). In the second round of screening, transgenic worms with *pmks-1::gfp* and *ppgp-12::gfp* reporters were treated with 50 mM EMS for mutagenesis. The *pgp-12* promoter driven GFP reporter was used as a control for screening. I expected to exclude mutants that affect regulators not working in neurons by introducing a second reporter *ppgp-12::gfp* in screens. I screened for the worms that had unchanged expression of p*pgp-12::gfp* and that had altered expression of *pmks-1::gfp*. I used same estimation to calculate screening genome size. The genome is covered for 94,200 times in second round of screen.

In order to eliminate heterozygote or low penetrance mutants, the mutants were maintained for three generations. Every time I maintained the worms, I picked up the mutant worms with altered GFP expression under a fluorescence dissecting microscope. After three generations, worms with stably altered GFP expression were selected. Worms from these three selected mutant lines were carefully observed under a confocal microscope (Table 3.2). Compared to wildtype, mutant worms in line 36-3-1, 11-3-1 have reduced GFP signal in inner labial neurons, amphids and phasdmis, but GFP signal in excretory cell is the same as widltype (Figure 3.13). The *pmks-1::gfp* expresses in same neurons in mutant line 36-3-1 and 11-3-1 and wildtype (Figure 3.14). *ppgp-12::gfp* expresses in the excretory cell (Figure 3.14). Mutants from line 39-1-7 have identical *pmks-1::gfp* expression in amphids and phasmids (Figure 3.13). Mutant line 39-1-7 have wildtype *ppgp-12::gfp* expresion in excretory cell (Figure 3.13). However, mutants from 39-1-7 have suppressed *pmks-1::gfp* expression in subset of inner labial neurons (Figure

3.13 and Figure 3.14). These mutants were crossed to N2 and wildtype *pmks-1::gfp* males respectively to determine whether they carry dominant mutations, recessive mutations or have defective transgenic arrays (Table 3.2). Mutants from line 36-3-1 and 11-3-1 have defective transgenic arrays. Mutants from line 39-1-7 have a recessive mutation on the X-chromosome. Mutant line 39-1-7 is named as JNC286 (*dot14*; *dotIs248 [pmks-1::gfp;dpy-5(+)]; dotIs296 [ppgp-12::gfp;dpy-5(+)]*). However, I did not avoid picking up mutants that have epigenetics modifications of the transgene with introducing *ppgp-12::gfp* reporter to screens. One possible reason is that two transgenes are of different loci. Thus, it is worth generating transgenic worms with two reporters in one array by co-injection for future genetic screens.

Table 3.2. Details of homozygous viable mutants selected from second round of genetic screening

Category	P0 plate numbers	Amount of P0 on plate	Mutant serial number	Mutant phenotype under dissecting microscope	Penetrance (observation of three generations)	Characteri-zation
Details	39	9	39-1-7 (JNC286)	*mks-1* expression pattern changes	100%	X-linked recessive mutations
	36	9	36-3-1	No *mks-1* expression	100%	Defective transgenic arrays
	11	5	11-3-1	Weak *mks-1* expression	100%	Defective transgenic arrays
	34	9	34-1-3	*mks-1* expression pattern changes	100%	Lost: all dead after three generation
	41	10	41-2-1	Weak *mks-1* expression	Low	Not studied
	43	10	43-1-2	Weak *mks-1* expression	Low	
	23	7	23-1-1	Weak *mks-1* expression	Low	

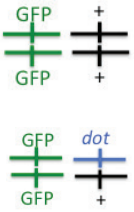
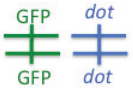
Note. The 39-1-7 was used for following study. The 39-1-7 was named JNC286.

After characterization, I found one homozygous recessive mutant, named *dot14* (strain JNC286). In cross of *dot14* mutant to N2, I observed that F1 outcrossed hermaphrodite progeny has wildtype GFP signal and F1 outcrossed male progeny has abolished GFP signal in subset of inner labial neurons. The F1 generation worms were

carefully observed under a confocal microscope. Outcrosses F1 progeny have different inner labial neurons expression in two genders. F1 male progeny have abolished GFP signal in subset of inner labial neurons, which is identical to *dot14* mutant GFP expression (Figure 3.15). F1 outcrossed hermaphrodite progeny have GFP signal in three pairs of inner labial neurons, which is identical to wildtype expression (Figure 3.15). Thus, *dot14* is an X-linked recessive mutation.

In order to confirm that *dot14* an X-linked recessive mutation, I crossed *dot14* mutant (strain JNC286) to wildtype *pmks-1::gfp;ppgp-12::gfp* transgenic worms (strain JNC285). I also observed that F1 outcrossed hermaphrodite progeny has wildtype GFP signal and F1 outcrossed male progeny has abolished GFP signal in subset of inner labial neurons. I count gene segregation ratio in F2 progeny. I scored phenotypes of the F2 progeny. I placed one L4-stage F1 generation worm in one plate. I set up two such plates. Every morning, I transferred the F1 generation worm to a new plate. I kept transferring F1 generation worms to new plates for four days, until the F1 generation worms were too old to lay more embryos. On the fourth day, I started scoring phenotypes of L4-stage or adult-stage F2 generation worms. After continuously scoring the F2 generation worms for four days, I added up number of same phenotypes. The F2 segregates either had abolished GFP signal in subset of inner labial neurons (mutant) or wiltype GFP signal in three pairs of inner labial neurons (wildtype), in a ratio of 1:3 (Table 3.3). In theory, from a heterozygous F1 with recessive mutant allele, 1/4 of the F2 segregates will be mutant, and 3/4 will be wildtype. Thus, I conclude that *dot14* mutant has a X-linked recessive mutation.

Table 3.3. Ratio of F2 progeny phenotypes in cross *dot14* mutant to wildtype *pmks-1::gfp* transgenic worms

Plate number	Category	Total	Phenotype	Genotypes	Phenotype	Genotypes	Ratio (Regular GFP: Mutant GFP)
			Regular GFP: GFP expression in labial neurons, amphids and phasmids		Mutant GFP: Only express in one labial neurons, normal expression in amphids and phasmids		
1	Number of F2 segregates from one F1 worm	359	270		89		3.03:1
2		327	258		69		3.74:1

Note. Each heterozygous F1 generation hermaphrodite of cross *dot14* mutant to wildtype *pmks-1::gfp* transgenic worms is self crossed on one plate. Phenotypes of F2 generation progeny on each plate are counted.

I carefully compared GFP expression pattern in *dot14* mutant and wildtype transgenic worms. Compared to wildtype, the *dot14* mutant lost expression in five inner labial neurons (Figure 3.16). In *dot14* mutant, the *pmks-1::gfp* expression is lost in ILL and ILV in both left and right asymmetrical side (Figure 3.17), yet it is unidentified whether these neurons are IL1 or IL2. In *dot14* mutant, the *pmks-1::gfp* expression is also lost in one of the two ILD, although it is not known whether it is lost in left or right side, and whether the *pmks-1::gfp* expressed neuron is IL1 or IL2 (Figure 3.17).

3.3.3. Preliminary genetic analysis of dot14 mutant (JNC286)

In order to know whether the *dot14* mutation in JNC286 strain affect labial neurons differentiation, I introduced *psrab-12::gfp* reporter to *dot14* mutant. I used strain BC12597, a transgenic strain that has GFP expression in labial neurons driven by *srab-12* promoter. In head region, *psrab-12::gfp* expresses in a set of neurons including inner labial neurons (Figure 3.18) (Hunt-Newbury *et al*. 2007). *srab-12* encodes a transmembrane serpentine receptor. Because *srab-12* is not a ciliary gene, it is not regulated by ciliary gene-specific transcription factors. If *dot14* affect neuron

development, *psrab-12::gfp* expression would be absent where inner labial neurons originally develop. If so, inner labial neuron is not differentiated properly in *dot14* mutant. The F1 male progeny from cross of *dot14* mutant to *psrab-12::gfp* wildtype transgenic worms has suppressed *pmks-1::gfp* expression in subset of inner labial neurons in accordance that *dot14* mutation is X-linked recessive. In *dot14* mutant, *pmks-1::gfp* expression is suppressed in four pairs of labial neurons (Figure 3.19). I observed F1 male progeny under a confocal microscope. I clearly observed GFP expression in three pairs of inner labial neurons in F1 male progeny (Figure 3.19). Thus, the labial neuron differentiation was not affected by the X-linked recessive mutation *dot14*. I conclude that *dot14* affect regulators of *mks-1* expression, but does not affect regulators of neuron development.

I would like to know whether the *dot14* mutation would affect other ciliary gene expression. The gene *tub-1* is a ciliary gene that express in subset of ciliated neurons. The gene *tub-1* encodes a TUBBY protein that regulates life span and fat storage pathways in *C. elegans* (Mukhopadhyay *et al*. 2005; Mukhopadhyay *et al*. 2007a). It is highly possible that *tub-1* and *mks-1* is regulated by a same group of transcription factors so that both genes express in subset of ciliated neurons, despite the fact that the only known ciliary gene transcription factor of *mks-1* and *tub-1* in *C. elegans*, DAF-19^{C} isoform, expresses in al ciliated neurons (Efimenko *et al*. 2005). I introduced *ptub-1::tdTomato* transgene to *dot14* mutant to test the hypothesis.

I used microinjection and X-ray integration to generate a transgenic strain of stable expressed *tub-1* promoter driven tdTomato (strain name JNC247). The *ptub-1::tdTomato* expresses in most ciliated neurons including amphids and phasmids (Figure 3.20). I chose tdTomato to track in-vivo *tub-1* expression so that I could distinguish it from *pmks-1::gfp* expression. The *ptub-1::tdTomato* transgene expression in strain JNC247 is identical with previous report (Efimenko *et al*. 2005). I used heat shock to introduce male strains carrying *tub-1* promoter driven tdTomato expression. I crossed the male *ptub-1::tdTomato* transgenic worms with *dot14* mutant hermaphrodites. The F1 hermaphrodite progeny is heterozygous in tdTomato, GFP and mutant allele loci. However, because the recessive *dot14* mutation is located on the X-Chromosome, the F1 male progeny is homozygous in *dot14* mutation loci. I observed F1 male progeny to know whether the X-linked recessive *dot14* mutation would affect *ptub-1::tdTomato*

expression. Confocal observation shows that in F1 male progeny, *ptub-1::tdTomato* expression is the same as wildtype transgene expression, while *pmks-1::gfp* expression is suppressed in inner labial neurons as expected (Figure 3.21). I conclude that *dot14* mutation does no affect *tub-1* expression.

Dyf phenotype (defect to take DiI or DiO) is an indication of *C. elegans* mutants that have malformation in cilia (Perens and Shaham 2005; Schafer *et al*. 2006). If a mutant have Dyf phenotype, I could conclude that cilia formation is abnormal in this mutant. However, some mutants with abnormal cilia, for example, ciliary gene i*fta-2* mutant still possesses a non-Dyf phenotype, and take DiI or DiO normally as wildtype worms in a dye-filling assay (Schafer *et al*. 2006). If the mutant has non-Dyf phenotype, I cannot make any conclusion on cilia formation. In this case, sub-cellular observation is required to understand cilia formation in the mutant (Williams *et al*. 2011).

I chose DiI for Dye-filling assay because mutants and wildtype carry GFP transgene, which is observable under blue fluorescence, while DiI is observable under red fluorescence. It reveals that 98% (60 out of 61 worms) of *dot14* mutant worms (JNC286 strain) take DiI in 6 pairs of amphids neurons and phasmids neurons. As control, 100% (90 out of 90 worms) wildtype transgenic worms (JNC285 strain) strain take DiI in 6 pairs of amphids neurons and phasmids neurons (Figure 3.22). Thus, *dot14* mutant has non-Dyf phenotype. The DiI went into 6 pairs of amphids neurons and phasmids neurons in both wildtype and *dot14* mutant (Figure 3.22). Thus, it is yet to be known that cilia formation in *dot14* mutant is abnormal or not.

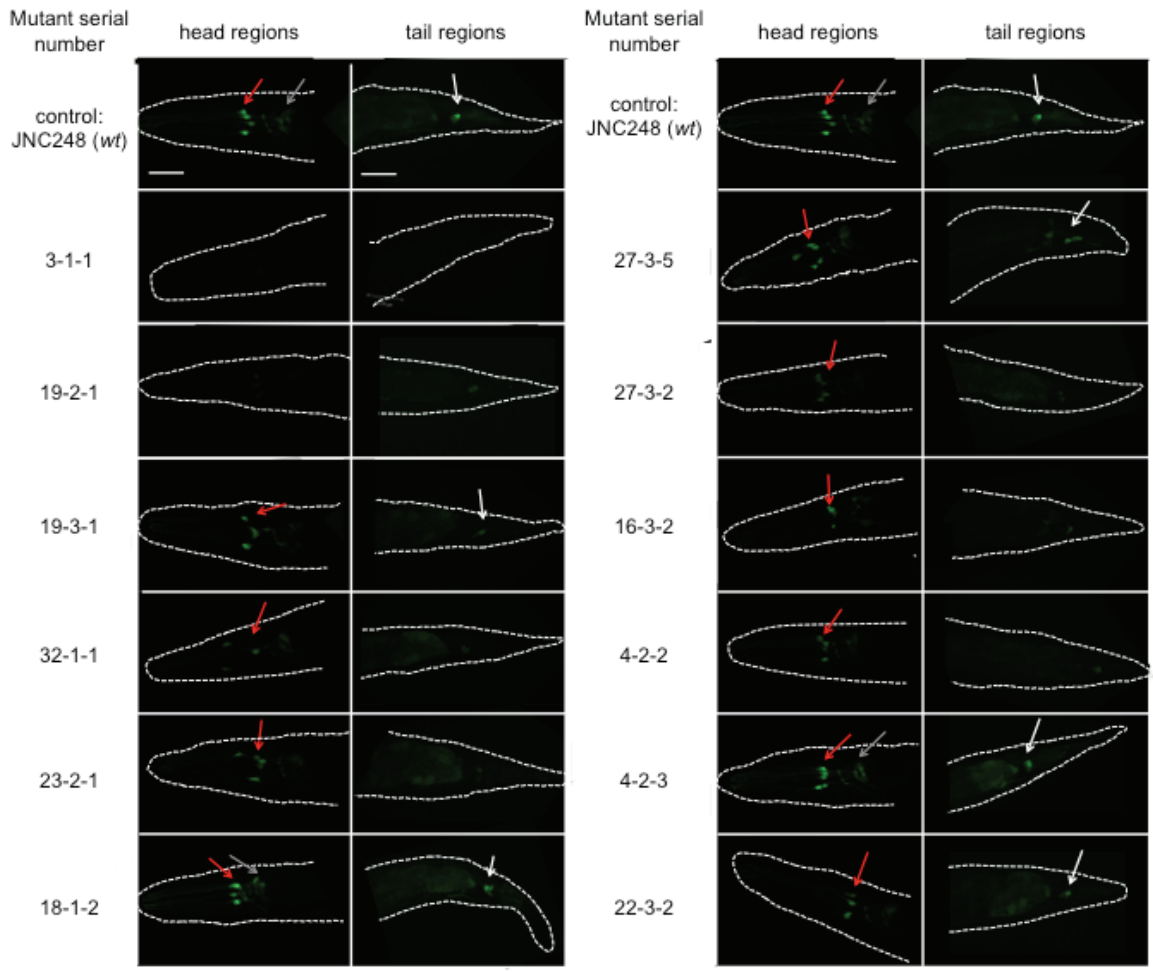


Figure 3.8. GFP signal levels in head and tail regions of 12 mutants

Note. Comparison of GFP signal strength between strating strain JNC248 with *mks-1* promoter driven GFP expression in wildtype background and mutant strains with *mks-1* promoter driven GFP expression in mutant background. First pane: wild-type p*mks-1::gfp* expression level in reporter strain. *pmks-1::gfp* express in labial neurons (red arrowheads), amphids (grey arrowheads) and phasmids (white arrowheads). Rest panels: *pmks-1::gfp* expression levels in mutant strains. The expression is significantly reduced in labial neurons, amphids and/or phasmids. Labial neurons were indicated by red arrowheads, amphids by grey arrowheads and phasmids by white arrowheads. The mutant worms in all images were observed immediately after isolation. All GFP exposure were taken at 345 ms. Contrast enhancement of all head images was set at black point 1750 and white point 44000. Contrast enhancement of tail images was set at black point 1750 and white point 10000. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 µm.

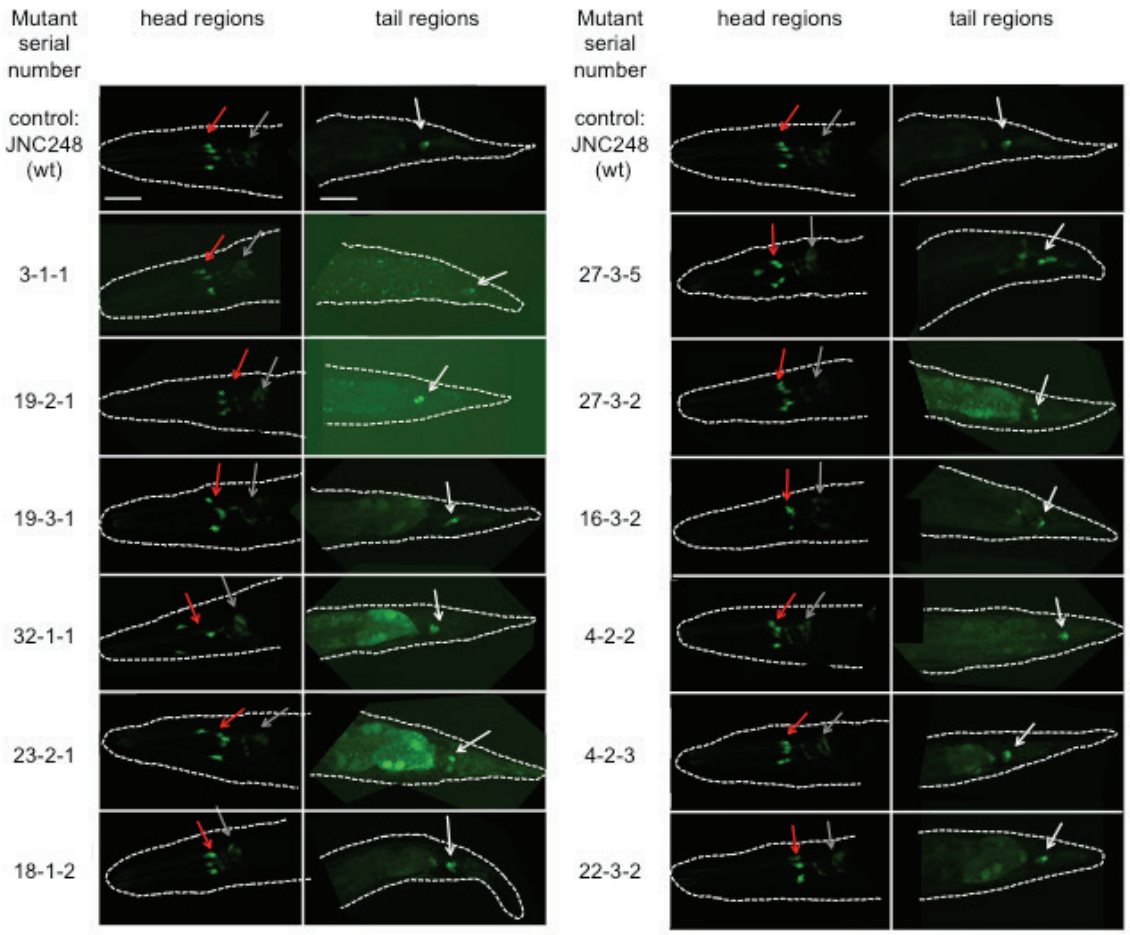


Figure 3.9. GFP signal patterns in head and tail regions of 12 mutants

Note. Comparison of GFP signal patterns between strating strain JNC248 with *mks-1* promoter driven GFP expression in wildtype background and mutant strains with *mks-1* promoter driven GFP expression in mutant background. First panel: wild-type p*mks-1::gfp* expression patterns in reporter strain. p*mks-1::gfp* expresses in labial neurons (red arrowheads), amphids (grey arrowheads) and phasmids (white arrowheads). Rest panels: *pmks-1::gfp* expression patterns in mutant strains. p*mks-1::gfp* expresses in labial neurons (red arrowheads), amphids (grey arrowheads) and phasmids (white arrowheads), which is similar to wildtype. All GFP exposure were taken at 345 ms. To display the GFP signal in all images, auto constrast enhancement were applied so that the signal would be visible. The mutant worms in all images were observed immediately after isolation. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 µm.

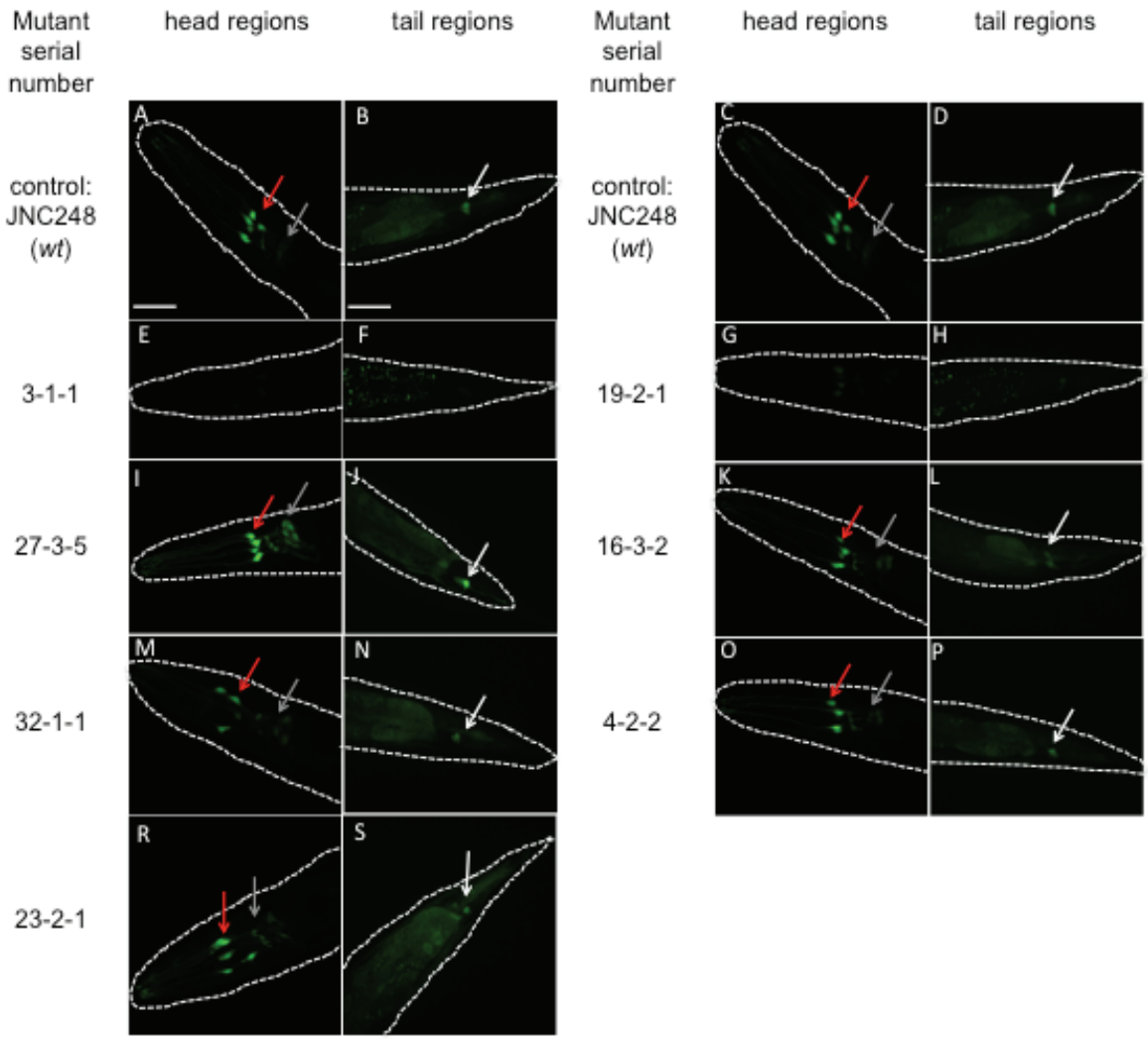


Figure 3.10. GFP signal levels in 7 mutants after 10 generations

Note. After maintaining the mutants for 10 generations, I observed GFP signal strength of 7 mutants using confocal microscope. A-D: wild-type p*mks-1::gfp* expression level in reporter strain. *pmks-1::gfp* expresses in labial neurons (red arrowheads), amphids (grey arrowheads) and phasmids (white arrowheads). E-H: *pmks-1::gfp* expression levels were significantly reduced in mutant strain 3-3-1 and 19-2-1. I-S: *pmks-1::gfp* expression levels in mutants that recovered to wildtype GFP expression levels after 10 generations. *pmks-1::gfp* expresses in labial neurons (red arrowheads), amphids (grey arrowheads) and phasmids (white arrowheads). These mutants that recovered to wildtype were discarded. All GFP exposure were taken at 354 ms. Contrast enhancement of all head images was set at black point 1750 and white point 44000. Contrast enhancement of tail images was set at black point 1750 and white point 10000. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 µm

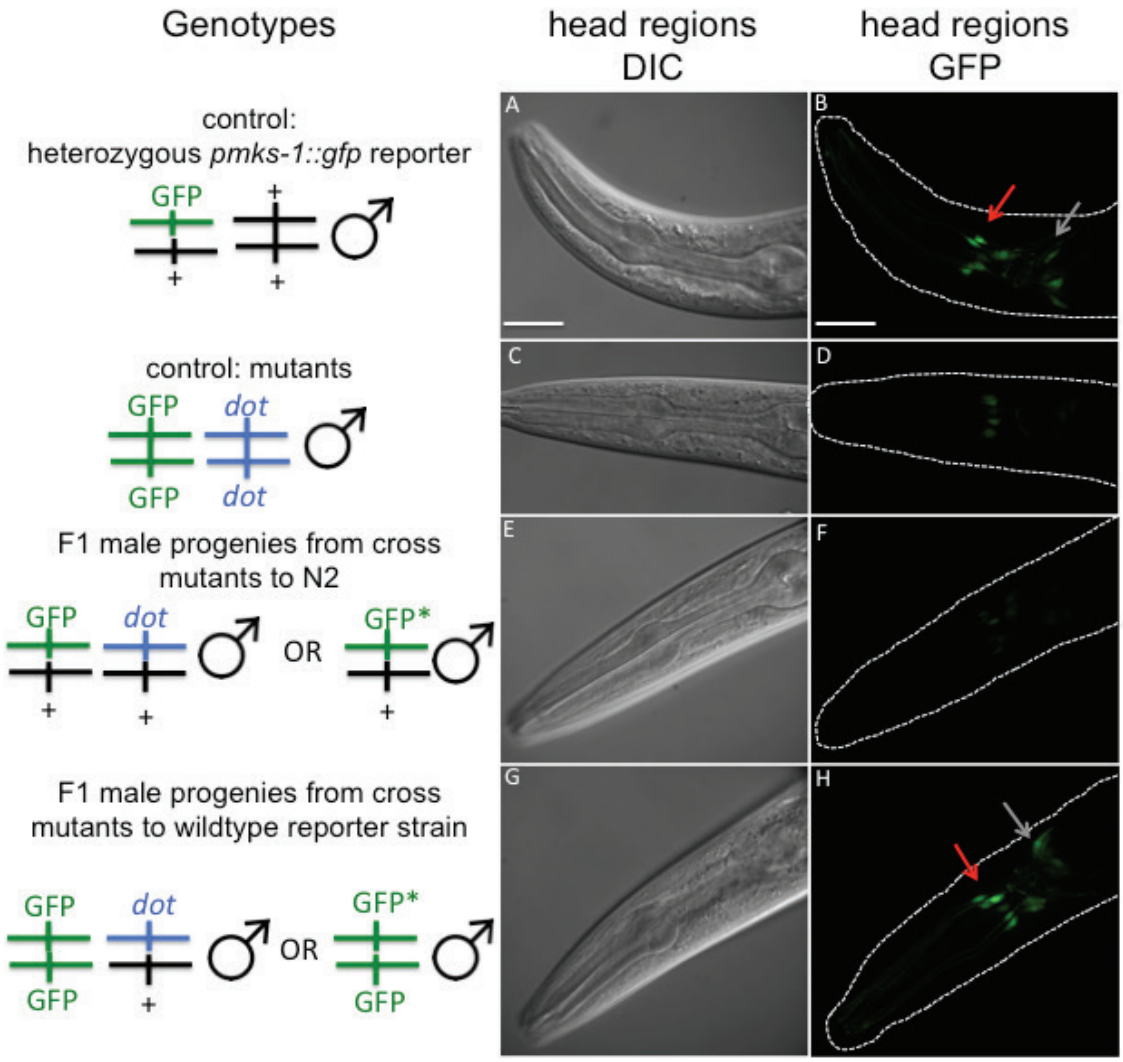


Figure 3.11. *dot6* mutant (strain JNC260) is characterized to carry defective transgene array

Note. Head regions were showed here because it is able to identify wildtype reporter expression and mutant expression by only showing head regions. A,B: wildtype *pmks-1::gfp* expression is the male worms that carries heterozygous reporter gene. The worms were male progeny from crossing *pmks-1::gfp* homozygous transgenic worms to N2. *pmks-1::gfp* expresses in labial neurons (red arrowheads), amphids (grey arrowheads). C,D: *pmks-1::gfp* expression in *dot6* mutant. Expression is significantly reduced in mutant worms. E, F: In F1 male progeny of cross mutant to N2, a copy of wildtype allele was introduced. F1 male progeny had significantly reduced reporter gene expressions. G, H: In F1 male progeny of cross mutant to wildtype reporter strain, an allele of intact transgene *pmks-1::gfp* copieswas introduced. F1 male progeny had wildtype reporter gene expression in labial neurons (red arrowheads), amphids (grey arrowheads). Thus, mutant strain JNC260 carries defective *pmks-1::gfp* transgenic arrays. GFP* stands for the transgene after EMS treatment. All GFP images were taken at exposure of 354 ms. Contrast enhancement was set at black point 2000 and white point 25000. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 µm

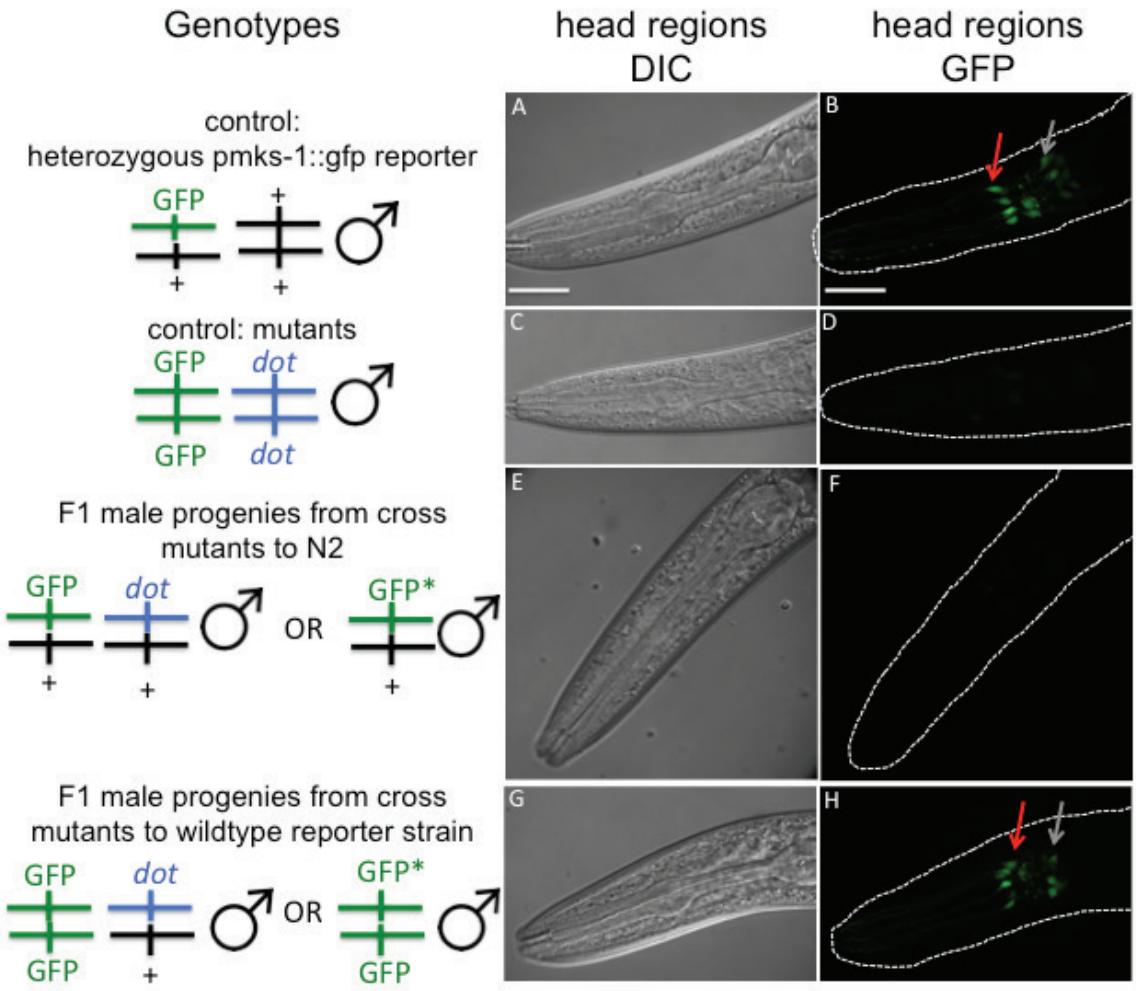


Figure 3.12. *dot7* mutant (strain JNC262) is characterized to carry defective transgene array

Note. Head regions were showed here because it is able to identify wildtype reporter expression and mutant expression by only showing head regions. A,B: wildtype *pmks-1::gfp* expression is the male worms that carries heterozygous reporter gene. The worms were male progeny from crossing *pmks-1::gfp* homozygous transgenic worms to N2. *pmks-1::gfp* expresses in labial neurons (red arrowheads), amphids (grey arrowheads). C,D: *pmks-1::gfp* expression in *dot7* mutant. Expression is significantly reduced in mutant worms. E, F: In F1 male progeny of cross mutant to N2, a copy of wildtype allele was introduced. F1 male progeny had significantly reduced reporter gene expressions. G, H: In F1 male progeny of cross mutant to wildtype reporter worms, an allele of intact transgene *pmks-1::gfp* copieswas introduced. F1 male progeny had wildtype reporter gene expression in labial neurons (red arrowheads), amphids (grey arrowheads). Thus, mutant strain JNC262 carries defective *pmks-1::gfp* transgenic arrays. GFP* stands for the transgene after EMS treatment. All GFP images were taken at exposure of 354 ms. Contrast enhancement was set at black point 1500 and white point 30000. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 μm

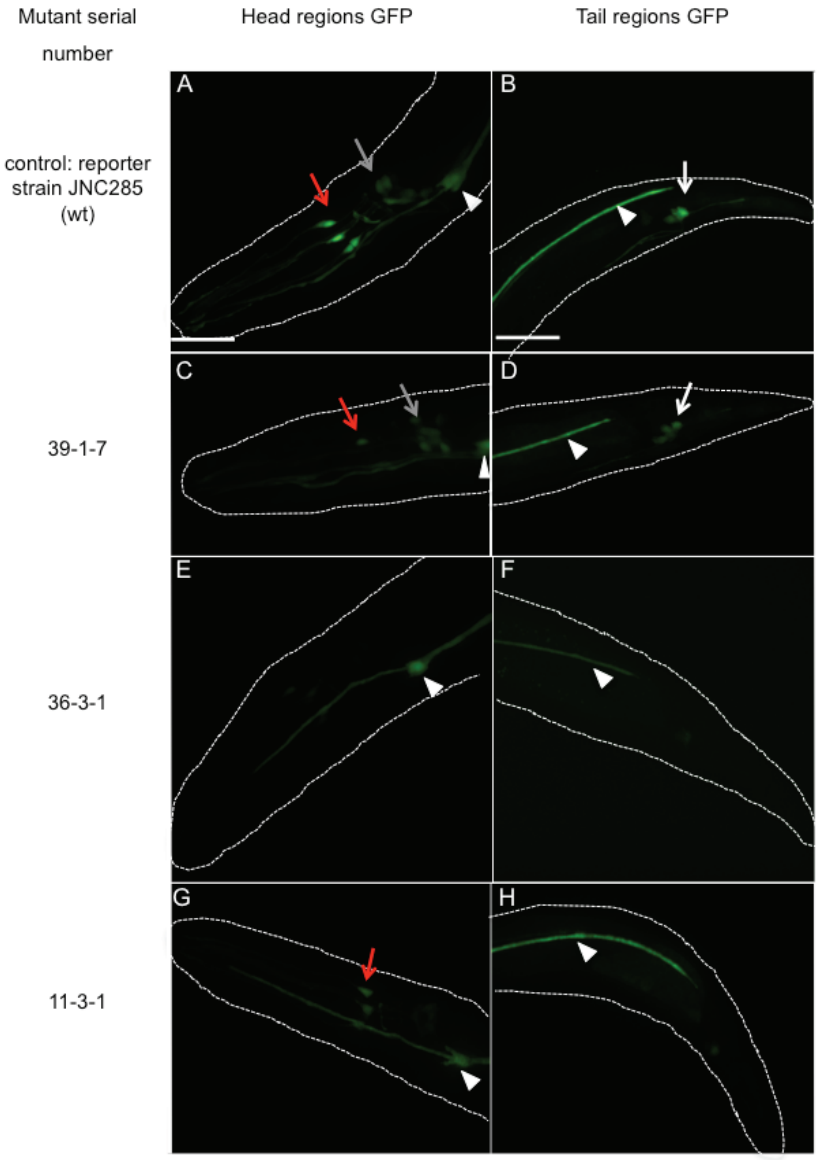


Figure 3.13. GFP expression levels in 3 viable mutants with100% penetrance from second round of genetic screening

Note. Comparison of *pmks-1::gfp* expression levels in wildtype reporter strain JNC285 and three mutants A, B: *pmks-1::gfp* expression level in wildtype reporter strain JNC285. *pmks-1::gfp* expresses in abial neurons (red arrowhead), amphids (grey arrowhead) and phasmids (white arrowhead). *ppgp-12::gfp* expresses in excretory cell (white triangles). C, D: *pmks-1::gfp* expression is suppressed in subset of labial neurons (red arrowhead) in mutant 39-1-7. *pmks-1::gfp* expresses in amphids (grey arrowhead) and phasmids (white arrowhead), similar to wildtype expression level. *ppgp-12::gfp* expresses in excretory cell (white triangles). E-H: *pmks-1::gfp* expression level is significantly reduced in these two mutants. Red arrowheads indicate labial neurons in image G. *ppgp-12::gfp* expresses in excretory cell (white triangles). All GFP exposures were taken at 200 ms. Contrast enhancement was set at head: black point 1700, white point 30000, tail: black point 1700, white point 10000. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 µm

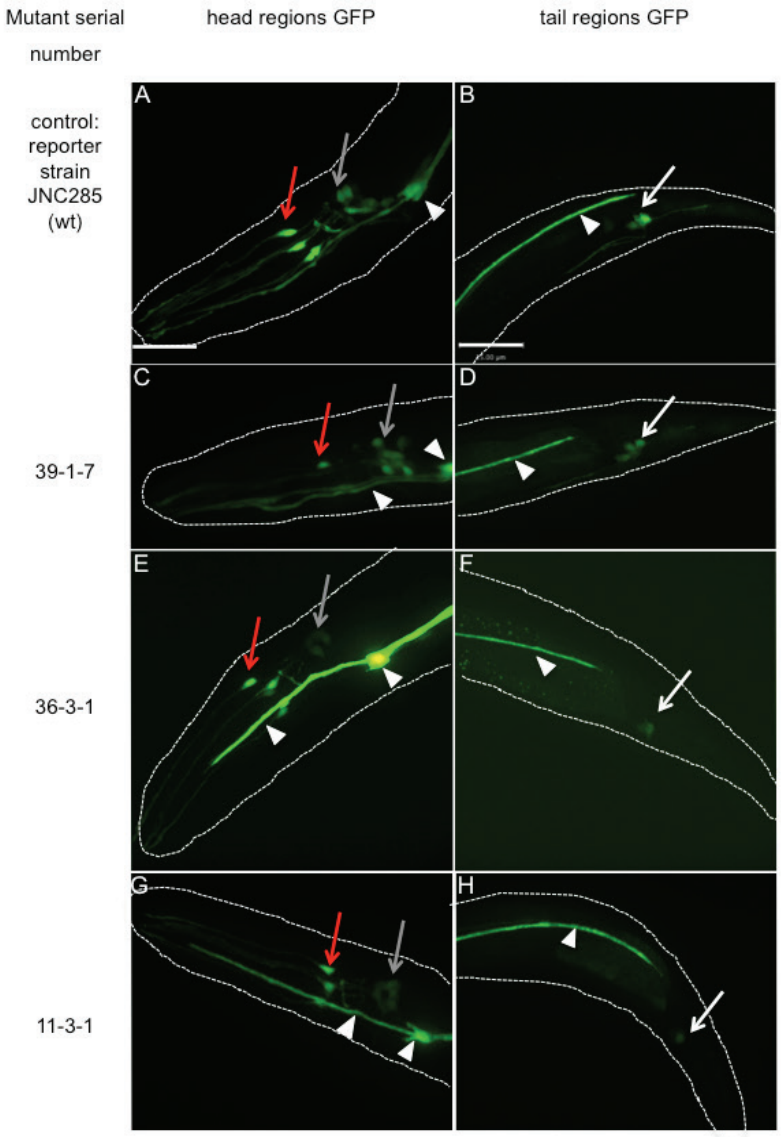


Figure 3.14. GFP expression patterns in 3 viable mutants with100% penetrance from second round of genetic screening

Note. Comparison of *pmks-1::gfp* expression patterns in wildtype reporter strain JNC285 and three mutants. A, B: *pmks-1::gfp* expression pattern in wildtype reporter strain JNC285. *pmks-1::gfp* expresses in labial neurons (red arrowhead), amphids (grey arrowhead) and phasmids (white arrowhead). *ppgp-12::gfp* expresses in excretory cell (white triangles). C, D: *pmks-1::gfp* expression pattern mutant 39-1-7. Expression is suppressed in subset of labial neurons (red arrowhead). *pmks-1::gfp* expresses in amphids (grey arrowhead) and phasmids (white arrowhead). *ppgp-12::gfp* expresses in excretory cell (white triangles). E-H: *pmks-1::gfp* expression pattern in two mutants. *pmks-1::gfp* expresses in inner labial neurons (red arrowhead), amphids (grey arrowhead) and phasmids (white arrowhead), smiliar to wildtype expression pattern. *ppgp-12::gfp* expresses in excretory cell (white triangles). All GFP exposures were taken at 200 ms. In order to display GFP signals, Autocontrast enhancement was applied to all images. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 µm

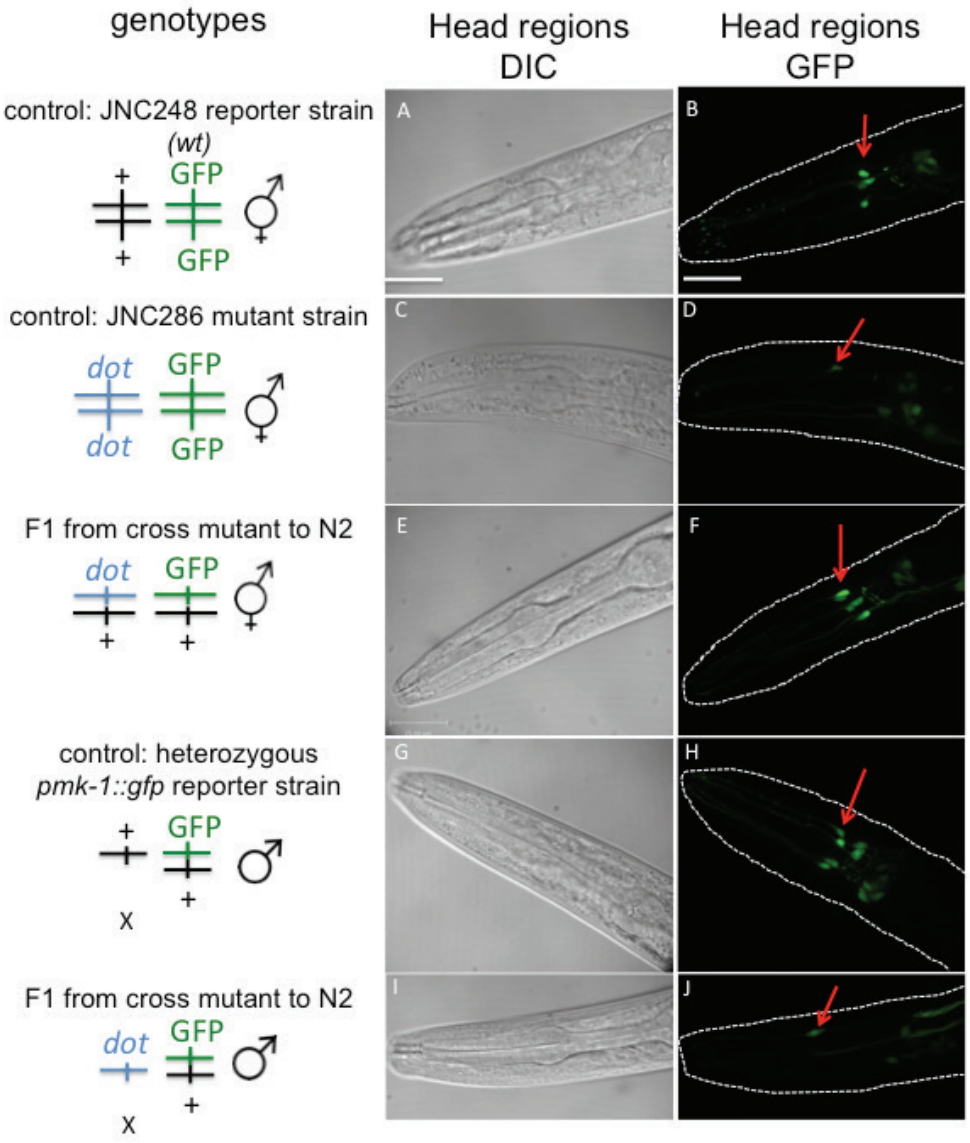


Figure 3.15. *dot14* mutant (strain JNC286) is characterized as X-linked recessive mutant.

Note. Head regions were showed here because it is able to identify wildtype reporter expression and mutant expression by only showing head regions. I crossed *dot14* mutant to N2, and observe F1 progeny. A, B, G, H: In wildtype strain, GFP signal is detected in three pairs of inner labial neurons (red arrowhead). C, D: In mutant strain, GFP signal is only detected in subset of inner labial neurons (red arrowhead). E, F: In F1 hermaphrodite progeny, GFP signal is detected in three pairs of inner labial neurons (red arrowhead), similar to wildtype. I, J: In F1 male progeny, GFP signal is detected in subset of inner labial neurons (red arrowhead), similar to *dot14* mutant. Thus, the mutant strain JNC286 carries X-linked recessive mutation that suppress *mks-1* expression in subset of inner labial neurons. All GFP exposures were taken at 131 ms. Contrast enhancement for GFP is: black point 1500, white point 40000. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 μm

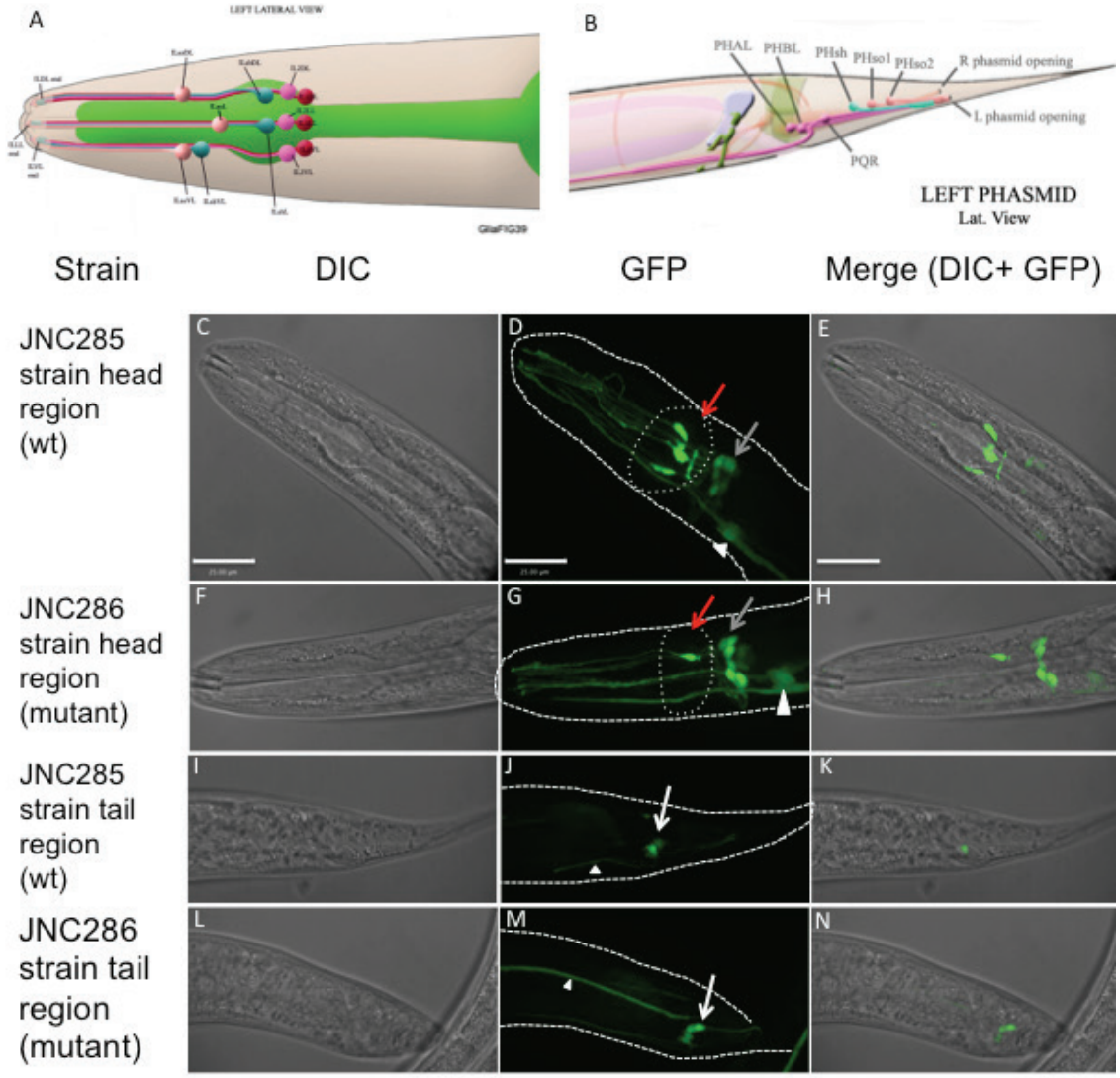


Figure 3.16 Expression of *pmks-1::gfp* in wildtype transgenic worm (strain JNC285) and *dot14* mutant (strain JNC286)

Note Comparison of *pmks-1::gfp* expression patterns widltype and mutant. A, B: Schematics of labial neurons and phasmids. Images obtained from wormatlas (Altun and Hall 2012). C-E: In wildtype reporter strain JNC285 head region, *pmks-1::gfp* expresses in inner labial neurons (dotted area pointed by red arrowhead), amphids (grey arrowhead). F-H: In *dot14* mutant strain JNC286 head region, mutation suppresses *pmks-1::gfp* expression in 2 pairs of inner labial neurons: ILLL, ILVL, ILLR, ILVR (dotted area pointed by red arrowhead), but *pmks-1::gfp* expression in amphids is similar to wildtype (grey arrowhead). Single layer images are also presented to identify inner labial neurons in wildtype transgenic worms and *dot14* mutant worms (see Figure 3.17). I-K: In wildtype reporter strain JNC285 tail region, *pmks-1::gfp* expresses in phasmids (white arrowhead). L-N: In mutant strain JNC286 tail region, *pmks-1::gfp* expresses in phasmids (white arrowhead), which is similar to wildtype expression. *ppgp-12::gfp* expresses in excretory cells (white triangles) in image C-N. All GFP images were at the exposure of 394 ms. All GFP images were confocal stacks of whole worm body from lateral view. Scale bar: 25 μm

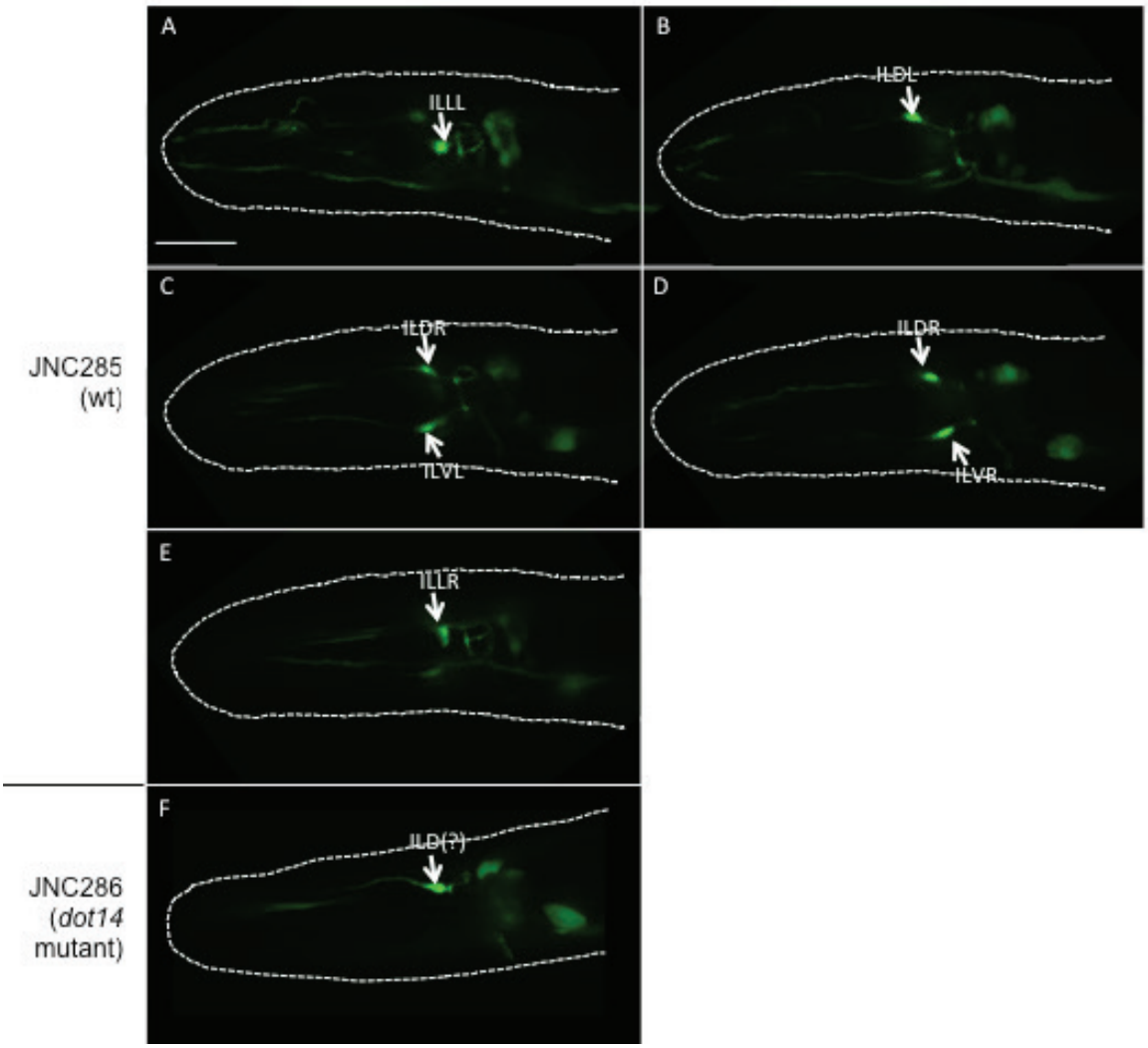


Figure 3.17 Expression of *pmks-1::gfp* in inner labial neurons in wildtype (JNC285) and *dot14* mutant (JNC286)

Note I captured single layer images of wildtype *pmks-1::gfp* transgenic worms (JNC285 strain) in head regions. 6 inner labial neuron cell bodies are observed. A-C: *pmks-1::gfp* expression is observed in ILDL, ILLL, ILVL, indicated by white arrowheads. D,E: *pmks-1::gfp* expression is observed in ILDR, ILLR, ILVR. It is unidentified whether these 6 cell bodies belong to IL1 or IL2 neurons. I captured single layer images of *dot-14* mutant (JNC286 strain) in head regions. One inner labial neuron cell body is observed. F: *pmks-1::gfp* expression is observed in ILD neuron, but it is hard to identified whether the neuron is in left or right side. It is also unidentified whether this cell body belongs to IL1 or IL2 neurons. Exposure time was 394 ms. Scale bar: 25 μm.

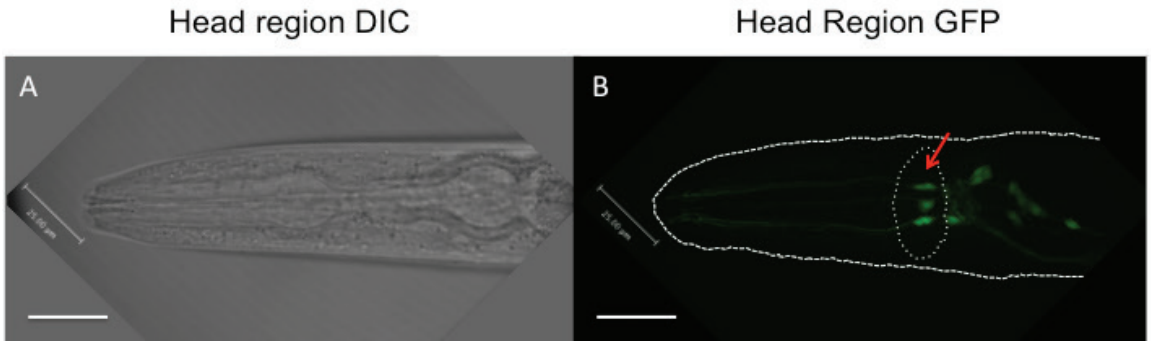


Figure 3.18. *srab-12* promoter driven GFP expression in labial neurons.

Note. A: DIC image of BC12597 (*psrab-12::gfp*) transgenic worm head region. B: *psrab-12::gfp* expression is detected in inner labial neurons (red arrowhead pointing dotted area). GFP exposure is 667 ms. GFP images are confocal stacks of whole worm body from lateral view. Scale bar: 25 µm. The strain was obtained from the Baillie lab.

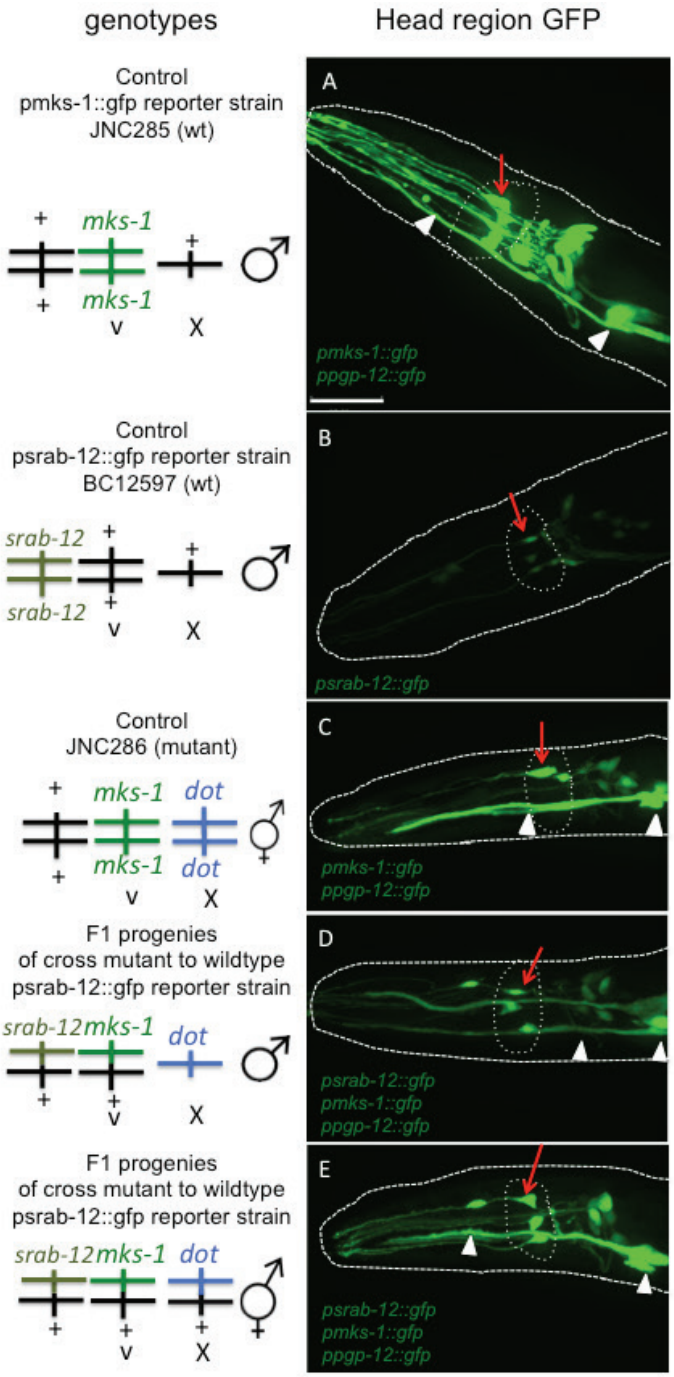


Figure 3.19. Observation of F1 progeny in cross *psrab-12::gfp* transgene (strain BC12597) to *dot14* mutant (strain JNC286)

Note. I crossed wildtype males that carrying *psrab-12::gfp* reporter to mutant hermaphrodites that carry X-linked recessive mutation (JNC286), and observed the F1 progeny males and hermaphrodites. Head regions were showed here because it is able to identify wildtype reporter expression and mutant expression by only showing head regions. A: In wildtype *pmks-1::gfp* transgenic worms (JNC285 strain), three pairs of inner labial neurons is visible with GFP signal (red arrowhead). B: In wildtype *psrab-12::gfp* transgenic worms (BC12597 strain), three pairs of inner labial neurons are visible with GFP signal (red arrowhead). C: In *dot14* mutant (JNC286), *pmks-1::gfp* expression is suppressed in subset of inner labial neurons (red arrowhead). D: In F1 generation males that carries *psrab-12::gfp* reporters, three pairs of inner labial neurons are visible with GFP signal. Although F1 generation males are X-linked recessive mutants, inner labial neuron differentiate normally in mutants. E: In F1 generation hermaphrodites, inner labial neurons are visible with GFP signal. Because F1 generation hermaphrodites only have one copy of recessive mutant allele, the phenotype is similar to wildtype. Thus, the conclusion is that *dot14* mutation does not affect inner labial neuron development. *ppgp-12::gfp* expresses in excretory cells in image A, C, D, E (white triangles). GFP exposure is 667 ms. GFP images are confocal stacks of whole worm body from lateral view. Scale bar: 25 μm.

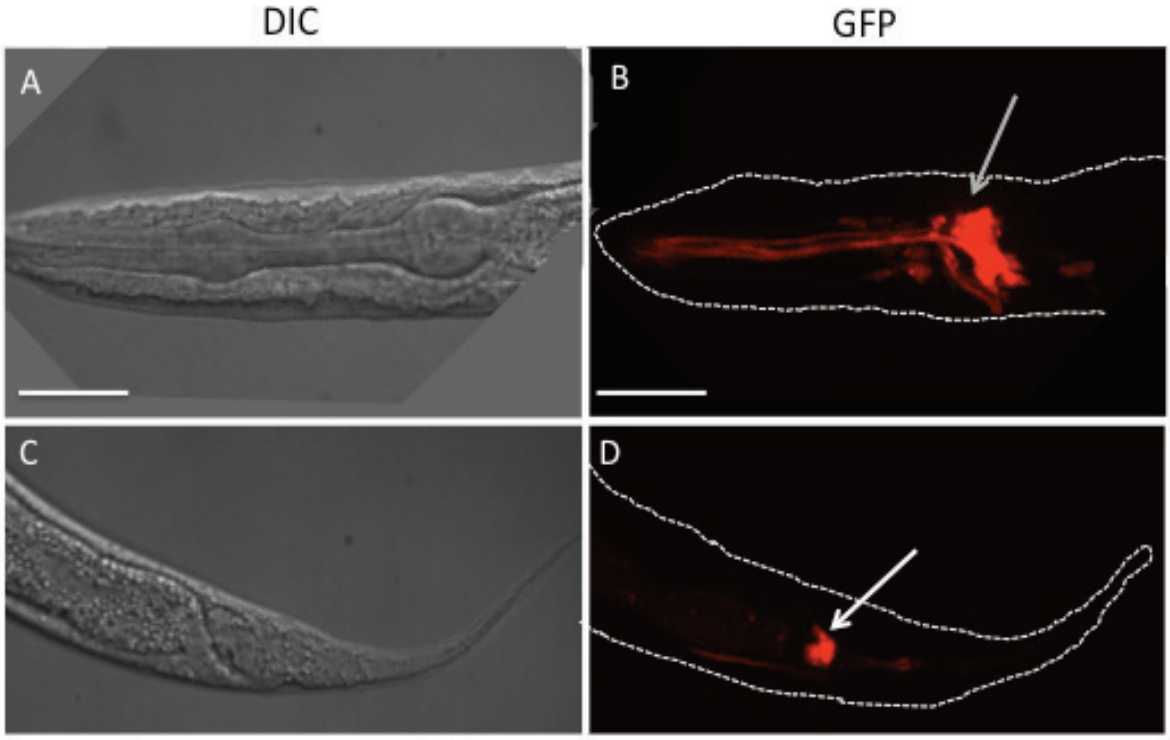


Figure 3.20. ***tub-1*** **promoter driven tdTomato expression pattern in JNC247 strain**

Note. A, B: In head regions, *ptub-1::tdTomato* expresses in amphids (grey arrowhead). C, D: In tail region, *ptub-1::tdTomato* expresses in phasmids (white arrowhead). RFP exposure is 111 ms. Autocontrast enhancement were applied to best display tdTomato signal. RFP images are confocal stacks of whole worm body from lateral view. Scale bar: 25 μm.

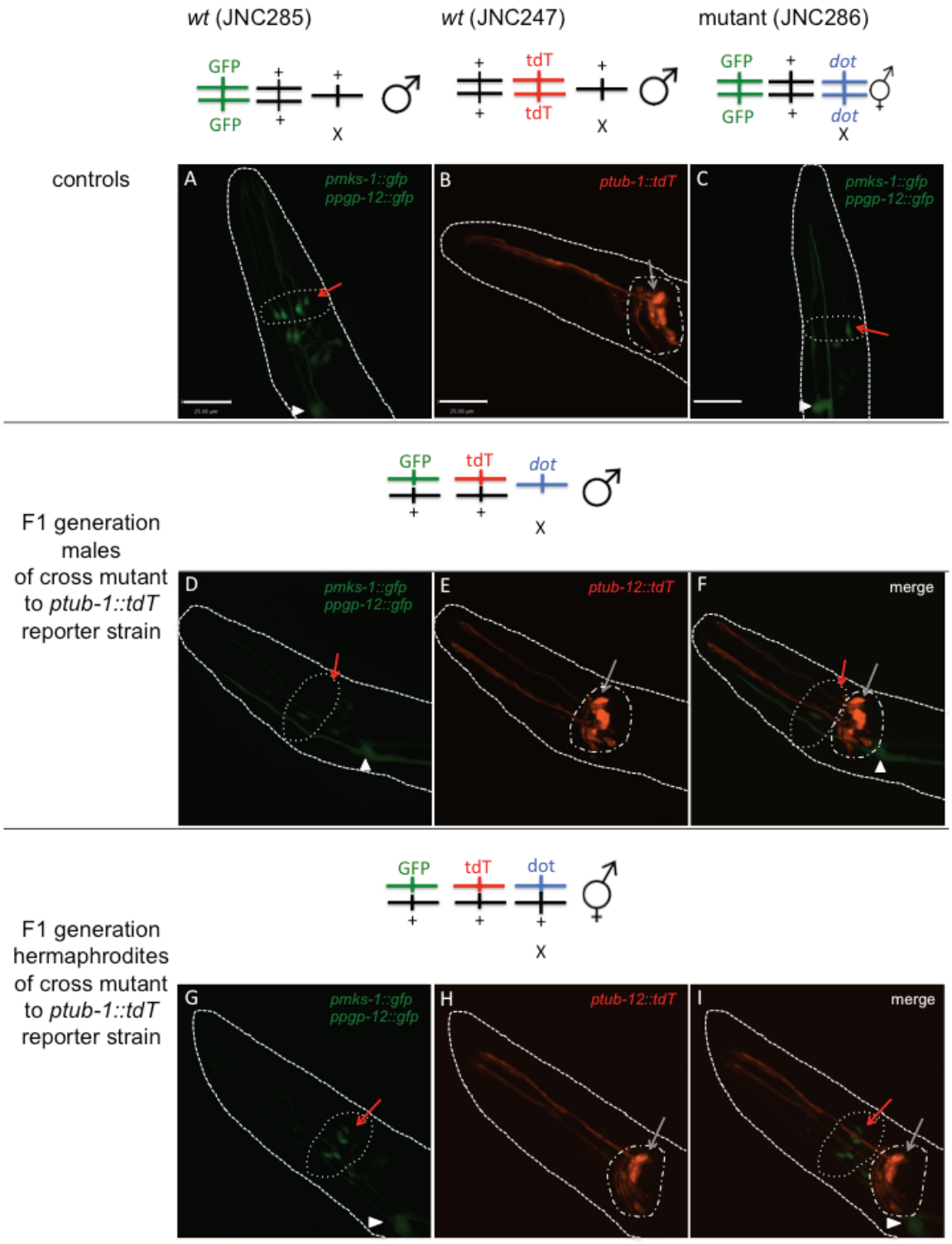


Figure 3.21. Observation of F1 progeny in cross *ptub-1::tdTomato* transgene (strain JNC247) to *dot14* mutant (strain JNC286)

Note. I crossed *ptub-1::tdTomato* reporter male strain (JNC247) to mutant hermaphrodites that carry X-linked recessive mutation (JNC286). I observed F1 progeny. Head regions were showed here because the mutations in JNC286 affect *pmks-1::gfp* expression in head region. A: In wildtype *pmks-1::gfp* transgenic worms, three pairs of inner labial neurons are visible with GFP signal (red arrowhead). B: In wiltype *ptub-1::tdTomato* transgenic worms, *ptub-1::tdTomato* expresses in amphids (grey arrowhead). C: In *dot14* mutant, mutations suppress *pmks-1::gfp* expression in subset of inner labial neurons (red arrowhead). D-F: In F1 male progeny from cross *ptub-1::tdTomato* transgenic worms to *dot14* mutant, mutation suppresses *pmks-1::gfp* expression in subset of inner labial neurons (red arrowhead). *ptub-1::tdTomato* expression in amphids (grey arrowhead) is similar to wildtype expression in image B. Because F1 generation males are X-linked recessive mutants, *dot14* mutation do not affect *ptub-1::tdTomato* expression. G-I: In F1 hermaphrodite progeny, *pmks-1::gfp* expression is visible in three pairs od inner labial neurons (red arrowhead), and *ptub-1::tdTomato* expression is visible in amphids (grey arrowhead). Because F1 generation hermaphrodites only have one copy of recessive mutant allele, the phenotype is similar to *wildtype ptub-1::tdTomato* and *pmks-1::gfp* reporter expression. Thus , the conclusion is that *dot14* mutation does not affect *ptub-1::tdTomato* expression. *ppgp-12::gfp* expresses in excretory cells in image A, C, D, F, G, I (white triangles). GFP exposure is 182 ms. RFP exposure is 111 ms. All images are confocal stacks of whole worm body from lateral view. Scale bar: 25 μm.

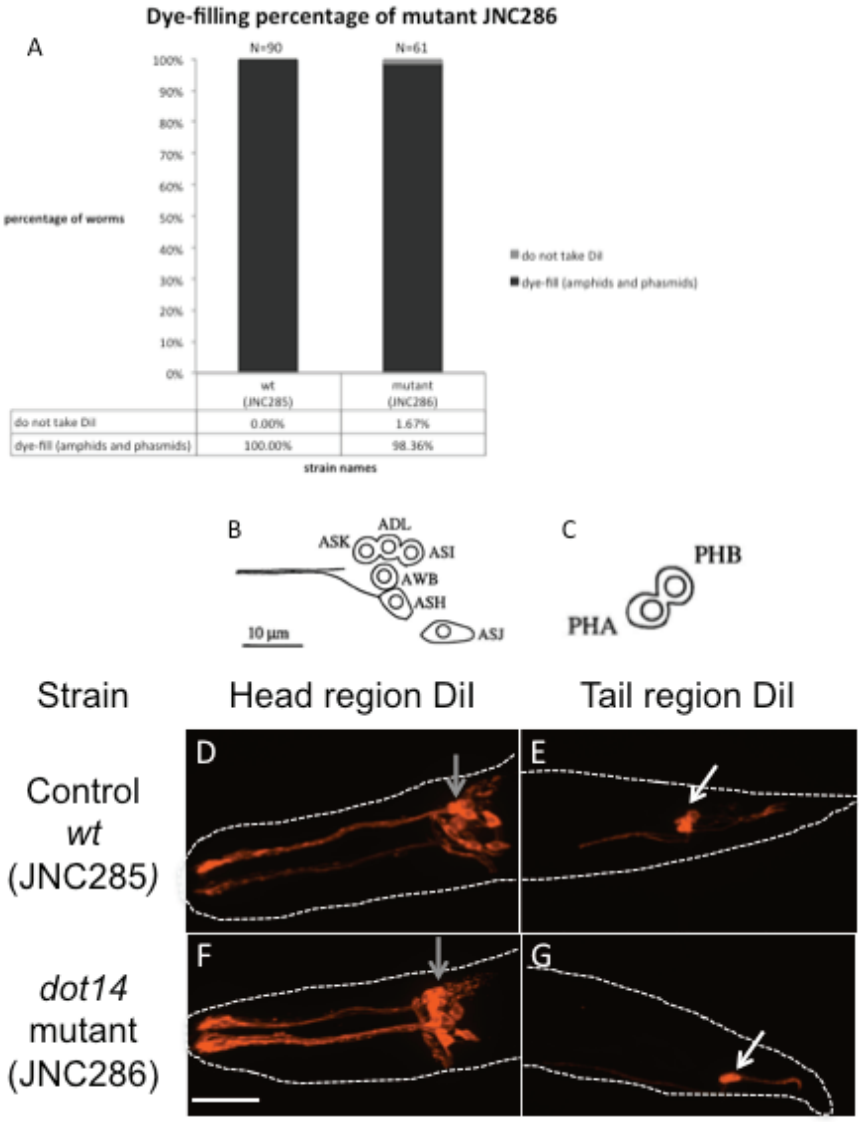


Figure 3.22 Dye-filling assay of *dot14* mutant (JNC286 strain)

Note. I scored the numbers of wildtype (JNC285) and *dot14* mutant (JNC286) that take DiI. I observed the DiI-filling patterns in wildtype (JNC285) and *dot14* mutant (JNC286). A: The percentages of mutant JNC286 that take DiI (98.36%), is almost the same as wildtype JNC285 (100%). B, C: The schematic diagram of 6 pairs of amphids neurons and two pairs of phasmids neurons that take DiI in wildtype worms. Images adopted from Starich, T.A. *et al.* (Starich *et al.* 1995). D, E: In wildtype worms (JNC285), DiI enters 6 pairs of amphids (grey arrowhead) and two pairs of phasmids (white arrowhead). F, G: The *dot14* mutant (JNC286) takes DiI in 6 pairs of amphids (grey arrowhead) and two pairs of phasmids (white arrowhead), which is the same as wildtype. RFP exposure is 91 ms. All images are confocal stacks of whole worm body from lateral view. Scale bar: 25 µm.

3.4. Conclusion

In my study, I applied EMS to introduce random mutations to wildtype worms genome that carries *pmks-1::gfp* as an indicator of screens. I used non-clonal F2 screens to look for mutants that have altered transgene expression. In non-clonal F2 screens, P0 worms laid all embryos. F2 progeny is a mixture of population from different F1 progeny. I screened all F2 progeny in first round, but I randomly screened 1/4 of F2 progeny in second round. This screening strategy avoided bias on slow-growing mutants. Another advantage of non-clonal F2 screen is that it saves time because there is no need to pick up individual F1 clones. However, it is not able to maintain lethal, sterile or dauer mutant alleles in a non-clonal F2 screens. In F1 clonal screens, after isolation of such mutant alleles, one can go back to the progeny of same F1 clones to maintain the mutation in a heterozygous genome (Rosenbluth *et al*. 1985). It should also be noted that the genome size for second round of screen might not be accurate. Although genome size could be estimated based on P0 worm numbers in non-clonal F2 screens, it is only accurate to the situation when all F2 progeny are screened. In second round of screen, only 1/4 of F2 progeny are screened. In another estimation, number of worms that one person could look at in a manual approach within one hour is recorded. And it is more accurate to use the calculation genome size= worm screened per hour (average) x total hours of screen x2.

In the screens, I expected to isolate mutants of candidate transcription factors that co-regulate *mks-1* expression with DAF-19. Multiple line of evidence reveal the possibility of existence of undiscovered transcription factors that specifically regulate *mks-1* expression or specifically regulate a group of ciliary genes expression (Efimenko *et al*. 2005; Mukhopadhyay *et al*. 2007b; Thomas *et al*. 2010; Choksi *et al*. 2014). After two rounds of genetic screens, I found one mutant, *dot14*, which carries a recessive mutation on the X chromosome. The *dot14* mutation affects suppresses expression of *mks-1* in subset of inner labial neurons. I applied simple genetic crosses to understand whether the *dot14* mutation non-specifically disrupts labial neuron development, and whether the *dot14* mutation affects other ciliary gene expression. The *dot14* mutation does not disrupt labial neuron differentiation because inner labial neurons are tracked by *psrab-12::gfp* expression in *dot14* mutant. The *dot14* mutation does not affect ciliary

gene *tub-1* expression, either. It is thus possible that *dot14* is located in a gene that specifically regulates *mks-1* expression in subset of ciliated neurons, which is correlated to my hypothesis. In regulation of ciliary gene *kap-1* expression in AWB neurons, such specific transcription factor exists. Cell-specific transcription factor FKH-2 functions downstream to DAF-19 to regulate *kap-1* expression in AWB neurons (Mukhopadhyay *et al*. 2007b).

Because EMS randomly introduces mutations to worm genome (Brenner 1974), I would expect mutants of multiple phenotypes. First, I expected that mutations were generated in a variety of genes that regulate worm development. For example, I isolated a vulvaless mutant in first screen. It is likely that vulva development is disrupted in this mutant. I isolated many sterile mutants. Mutations in genes that play a role in germline development might cause sterile phenotypes. Some of the isolated mutants are larva-arrest. Mutations in genes affecting life span might cause larva-arrest phenotype. All these vulvaless, sterile and larva-arrest mutants might carry mutations in the regulators upstream of *mks-1* specific regulation pathway, so these mutants also have reduced *pmks-1::gfp* expression. In second screen, I used *ppgp-12::gfp* transgene as a control so that I did not pick up vulvaless, sterile or larva-arrest mutants. Second, I expected to see mutants with Unc phenotype. There are 112 members of *unc* gene family. Based on EMS mutation rate and screening genome size, at least one mutation would hit an *unc* gene. I saw *unc* mutants during screen as expected. Because *daf-19* is a known transcription factor for *mks-1*, I also expected to isolate *daf-19* mutant. *daf-19* null mutation cause constitutively dauer-formation (Swoboda *et al*. 2000). I isolated some dauer mutants but was not able to maintain and characterize them. Probably these dauer mutants are *daf-19* null-mutant. Due to disadvantage of non-clonal F2 screen, I was not able to keep any dauer mutant.

In this study, I improved screening methods. In second round, I avoid wasting time on sterile, larva-arrest or vulvaless mutants because of using a second reporter *ppgp-12::gfp*. I improved selection of mutants by selecting one mutant from same P0 plate because multiple mutants from same P0 plate are likely to be segregated from same F1 mutant progeny. I observed mutant penetrance for three generations to select mutants with complete penetrance. I applied genetic cross to characterize mutants as recessive, dominant or transgene defective. It is not clearly understood what results in

the defective transgenic arrays in screens. Doitsidou *et al.* proposed that transgene is silenced by mutations in transgene array (Doitsidou *et al*. 2008). In my study, I used a transgenic strain with multiple copies of the *pmks-1::gfp* transgene. Based on estimation, there are 20 copies of *pmks-1::gfp* transgene in the worms. Thus it is hard for the EMS mutagen to hit all copies of transgene. Two possibilities could cause transgenic-array defects in my study. First, the copy number of transgene array is reduced because non-homologous recombination happens. Second, chromatin enzymes silence the transgene expression at an epigenetics level by histone modifications. I introduced a second reporter *ppgp-12::gfp* to avoid picking up transgenic-array defective worms, but I still got such worms in second screen. Two reporters, *pmks-1::gfp* and *ppgp-12::gfp,* are located on different loci. Thus, it is important to generate transgenic worms that carry both reporters integrated to same locus on the chromosome by co-injection. In this case, both transgene reporters' expression would be suppressed if chromatin enzyme modifies same location of histone. For further genetic screens, it should be considered to generate transgenic worms by co-injecting two reporter constructs.

I identified a recessive *dot14* mutation on the X chromosome. The *dot14* mutation is likely to affect regulators of *mks-1* expression in inner labial neurons. Whole genome sequencing and SNP-mapping is required to identify the location of the *dot14* mutation, the nucleotide information of the mutation, and the affected gene (Doitsidou *et al*. 2010). To prepare for whole genome sequencing, I cleaned the *dot14* mutant (strain JNC286) by backcrossing it four times to wildtype transgenic worms (strain JNC285). The backcrossed strain was named JNC319 (*dot14*; *dotIs248 [pmks-1::gfp;dpy-5(+)]; dotIs296 [ppgp-12::gfp;dpy-5(+)]*). In the future, colleagues in Chen lab will sequence the whole genome of the cleaning mutant strain JNC319, and use bioinformatics tools to map the *dot14* mutation. Based on genetic features of the mutation, which I provided in my thesis, lab colleagues can collect bioinformatics data to identify the mutated gene, which plays important role in transcriptional regulation of *mks-1*.

Chapter 4.

General Conclusion

4.1. Conclusion and Discussion

I used forward genetic screens to isolate mutants of genes that encoding transcription factor regulating *mks-1* expression. Genome-wide mutagenesis is applied in forward screens to search for unknown genes controlling a biological pathway. Forward genetic screens have been applied to identify transcription factors in many studies. For example, Mango *et al*. isolated mutant alleles of a transcription factor encoding gene *pha-4* in a screen to isolated first stage larva lethal mutant (Mango *et al*. 1994). Homozygous *pha-4* mutant arrests at embryos or first stage larva. Microscope observation shows that *pha-4* mutant lacks pharynx. They are able to maintain the recessive mutant allele in a heterozygous form, which appears as wildtype phenotype. *pha-4* encodes PHA-4 protein that regulates proper pharyngeal cell differentiation. Later on, it is found that PHA-4 encodes a FOXA transcription factor in early embryonic stage (Kiefer *et al*. 2007). In Mango *et al.* study, they used 0.5 M ethylnitrosourea or 0.05 M methanesulfonate to conduct several rounds of screens. In total they obtained seven recessive alleles of *pha-4* gene. They used clonal screen so they could maintain lethal alleles in a heterozygous form after isolation. Mango *et al*. searched for a lethal phenotype, which could be seen among mutants that affect any regulators of worm development. Compared to their study, I introduced reporter transgene to track target gene expression. I monitor target gene expression to search for mutants that affect target gene expression. The mutant pool is small and specific on *mks-1* regulation. The reason to introduce transgene to monitor mutant phenotype is that target gene *mks-1* only expresses in subset of ciliated neurons. Thus, defects in transcription factor that specifically regulate *mks-1* expression might merely cause changes of *mks-1* expression in subset of cells instead of causing severe functional malformations. The subtle change

in expression is visible with GFP transgene expression. I use non-clonal screen in my thesis study, because I did not expect to pick up and maintain mutants with severe developing problem. This kind of mutant is more likely to hit transcription factors, such as DAF-19, that regulate all ciliary gene expression or worm development.

Introducing GFP transgene to monitor gene expression in genetic screens is a mature technique. Compared with other similar GFP genetic screens, I would like to discuss some interesting phenomena. First I found that some mutants are actually transgenic worms with defective transgenic arrays. Doitsidu *et al*. also find this kind of mutants in a screen for mutations affect dopaminergic neuron differentiation (Doitsidou *et al*. 2008). They used GFP transgene that contains dopaminergic neuron specific gene *dat-1* fused to GFP. They identified mutants that carry defective transgenic arrays, by using same genetic cross that were described in my thesis. Although they find that mutants have mutations in transgenic arrays (Doitsidou *et al*. 2008), I propose other possibilities that generate defective transgenic arrays. The transgenic worms carries at least 20 copies of transgenic constructs (Dr. D. Baillie, personal communication), and thus it is theoretically hard to generate mutations in all copies of transgene array to silence transgene expression (Mello and Fire 1995). My lab colleagues are using same method to screen for mutants that have other ciliary gene expression changes. They also found “mutants” that actually had defects in transgenic arrays. There could be two hypotheses: a chromatin silencing due to histone modification at an epigenetics level (Kelly and Fire 1998), or reduction of transgene copy numbers due to non-homologous recombination (J. Wang, personal communication). Histone is modified by methylation, acetylation, phosphorylation and ubiquitination in some locations, thus making DNA unapproachable by transcription complex (Bannister and Kouzarides 2011). It is possible that some transgene expression is silenced by chromatin-enzyme induced histone modification near transgene copies. Based on Doistsidou *et al*. report and result in this thesis, transgene defective mutants occur in two independent screens. It is worth discussing whether defective transgenic arrays occur naturally in a certain proportion of transgenic worm population. If the rate could be determined, we can decide screening genome size for an efficient screen based on the rate. Second, “worm sorter” (COPAS Biosort machine) has been widely used to aid genetic screens. The worm sorter is a flow cytometry instrument to assort live worms based on their fluorescence density or size

(Pulak 2006). Compared with using manual screens, using the worm sorter saves time. In Doitsidou *et al*. study, they compared speed of manual screens and applying the worm sorter to screen (Doitsidou *et al*. 2008). The speed increased at least 10 times after applying the worm sorter. 10-laborous day manual work was finished within 1.5 day in a part-time manner. Meanwhile, the efficiency of the worm sorter approach is higher than the manual approach. Within same time, Doitsidou *et al*. obtained alleles of six loci by the worm sorter approach, while they only isolated alleles of two loci by the manual approach. The reason is likely to be that more samples were screened in the worm sorter approach within same time period. The sensitivity of the worm sorter is higher than human eyes. In the same study, mutants with as little as one cell losing GFP signal were retrieved even though the wildtype transgenic worms have GFP signal in eight different cells in head and tail. I used a manual approach due to limited machine availability in our lab. However, I would recommend using a worm sorter if it is available. Human eyes are not as accurate as high-fidelity cameras in detecting signals. With application of worm sorters, it will be possible to detect slightly altered signals (Stoeckius *et al.* 2009; Fernandez *et al.* 2010).

I have to admit that some limitations exist within this thesis work. First, I was not able to study any regulation of intronic region or downstream regions. I used a 2.1 kb upstream intergenic region fused to GFP to drive transgene expression. Although this upstream intergenic region contains minimal set of cis-regulatory sequences that is sufficient to regulate gene expression, cis-regulatory sequences are discovered in introns and downstream sequences (Okkema *et al*. 1993; Wagmaister *et al*. 2006). In my thesis study, I focused on regulators that bind to upstream sequences of *mks-1*. Second, I was not able to isolate any lethal or dauer mutant by using a non-clonal F2 screens. The *C. elegans mks-1* hypomorphic mutant allele *tm2705* does not cause lethal or dauer phenotype (Williams *et al*. 2008). If candidate regulators control *mks-1* expression level or gene “on”/”off” state in specific neurons, the mutants are not likely to carry lethal phenotype. Non-clonal F2 screens enable me to focus on mutants that affect this kind of candidate regulators. But I could not detect mutants of transcription factors that regulate all ciliary gene expression or that regulate neuron development pathways, for example, DAF-19. Third, I was not able to detect repressors that regulate *mks-1* expression. Because *dot14* mutation is recessive on X chromosome, it is likely to be a loss-of-

function mutation in a gene encoding an activator in *mks-1* regulation. In my thesis work, I screened for reduced GFP signal only. Mutants with reduced GFP signal might carry three kinds of mutations: a recessive loss-of-function mutation in a gene encoding an activator, a dominant loss-of-function mutation in a gene encoding an activator or a dominant gain-of-function mutation in a gene encoding a repressor. A dominant/recessive loss-of-function mutation on repressor gene results in increased GFP signal in mutant worms. The phenotypic criteria on selecting reduced GFP signal limited the possibilities to detect more repressors in *mks-1* regulation. Moreover, the amount of mutants isolated from this thesis work is smaller than expected. For example, Doitstidou *et al*. screened genome for 11,000 times manually, and they identified 10 mutants (Doitsidou *et al*. 2008). After isolation, they identified mutants based on similar criteria this thesis work used. They only picked up independent isolates from different P0 plates or have different phenotypes if from same P0 plate. However, in this thesis work, I merely identified 22 mutants for a genome overage of 40,200 times. It is possible that phenotypic screening criteria in my thesis work were too strict. Mutants with increased GFP were not picked up. Those mutations might be loss-of function of genes encoding repressors. It is possible that *mks-1* expression is suppressed by repressors in subset of ciliated neurons.

Other approach could be applied to identify transcription factors that regulate *mks-1* expression. Yeast One Hybrid Assay could be applied to identify proteins that bind to *mks-1* upstream/ downstream intergenic regions and intronic regions *in vitro*. Y1H assay methods for transcriptional regulation study in *C. elegans* is reviewed by Reece-Hoyces and Walhout (Reece-Hoyes and Marian Walhout 2012). The interested DNA sequences (e.g. *mks-1* upstream/ downstream intergenic regions and intronic region) are connected to a reporter gene that generates one kind of amino acid in yeast. The Y1H library includes transcription factors identified in *C. elegans* (Reece-Hoyes *et al*. 2005). Library proteins are fused with activation domain that activates reporter gene expression. The expression construct (library proteins fused to activation domain) and the bait reporter construct (DNA sequences fused to reporter gene) are transformed into yeast mutant that is not able to grow on selective medium that lacks the reporter-gene encoding amino acid. If interaction exists between a library protein and DNA sequences, reporter gene is activated so that yeast clone will grow on the selective medium. Y1H

assay can only detect direct interaction between transcription factor and DNA, but it is not applicable to study transcription factor in a pathway that regulate gene expression by targeting other direct-binding transcription factor.

I isolated recessive X-linked *dot14* mutation that specifically suppresses *mks-1* expression in five inner labial neurons. In *dot14* mutant, inner labial neuron differentiates and develops properly. Based on results from my thesis work and other studies, I propose two models for *mks-1* regulation in *C. elegans*. In first model, an unknown transcription factor recruits DAF-19 to *mks-1* regulatory sequences in a co-regulation manner to activate *mks-1* expression in specific ciliated neurons. Evidence of such co-regulation manner is reported as RFX3 and Fd3F regulate ciliary gene expression in motile cilia of *Drosophila* (Newton *et al*. 2012). In this case, it is interesting to know whether the unknown transcription factor physically interacts with DAF-19 or binds to *mks-1* promoter directly to co-regulate *mks-1* expression. In second model, an unknown transcription factor acts upstream or downstream of DAF-19 to control DAF-19 activity in specific ciliated neurons. DAF-19 is activated by the unknown transcription factor in subset of ciliated neurons in this model. Such regulation is reported in *kap-1* regulation in AWB neuron. It is reported that forkhead transcription factor FKH-2 regulated *kap-1* expression specifically in AWB neuron (Mukhopadhyay *et al*. 2007b). Many studies revealed that MKS-1 participate in activities of docking centriole to membrane to form basal body/transition zone in specific ciliated cells (Dawe *et al*. 2007; Weatherbee *et al*. 2009; Cui *et al*. 2011). My study shows that *dot14* mutation does not affect ciliary gene *tub-1* expression in amphids and phasmids. It would be interesting to know whether the unknown transcription factor is a gene-specific regulator for *mks-1* expression, or it is a cell-specific regulator for a group of ciliary gene regulation in inner labial neurons.

The methodology in my thesis work could be applied to studies of different ciliary genes on identifying gene-specific or cell-specific transcription factors.

4.2. Future Work

The next step is to analysis the mutant strain with Bioinformatics tools. I prepared genomic DNA of *dot14* mutant (JNC319 strain, 4x backcrossed strain from mutant

JNC286) and wildtype *pmks-1::gfp* transgenic worms (JNC285 strain). My lab colleagues will sequence the whole genome of the *dot14* mutant and the wildtype transgenic worms. They will map the mutations using SNP-mapping based on standard method (Doitsidou *et al*. 2010). They will use comparative genomics tools to identify genomic variations in the mutant strain. They will identify candidate genes that are affected by the variations. They will use candidate gene approach protocol (see Appendix) to genetically analyze candidates. Then they can identify the mutated gene that encodes transcription factors in *mks-1* regulation.

After identifying the unknown gene that *dot14* mutations hit, it is necessary to understand protein encoded by this gene, most likely a transcription factor, in two aspects: structure including DNA binding characteristics and function including roles in different pathways. First, it is important to understand the structure of transcription factor: either it has a domain of zinc finger, forkhead, winged-helix, homeobox or other unknown functional domain. Also, it is interesting to know the type of *dot14* mutation: is it a null-mutation or a hypomorphic mutation? It is interesting to know the transcription target site on a genome-wide scale. Chromatin Immunoprecipitation (ChIP) assay can be applied to identify the target sites. Many transcription factor target sites have been identified by ChIP assays in the modENOCODE project e.g. (Niu *et al*. 2011). It is wise to search wormbase for available data of target sites before any experiments. Second, it is necessary to test both models proposed in my thesis. To test first model, my colleagues could apply Co-immunoprecipitation (Co-IP) to test direct physical interactions between DAF-19 and the identified transcription factor. They could use DNA footprinting to detect direct target sites for identified transcription factor in *mks-1* promoter sequences. To test second, model my colleagues could use RNAi to interfere transcription factor expression. By RNAi of identified transcription factor, they can track *daf-19* expression using GFP fused transgene, or they can label *daf-19* mRNA product by *in situ* hybridization. They can place identified transcription factor in a signalling pathway that regulate specific gene expression or regulate ciliary gene expression in specific cells. Moreover, multiple genetic crosses could be done to answer the question whether identified transcription factor is cell-specific or gene-specific. My colleagues can introduce reporter transgene of different ciliary components into *dot14* mutant to track changes in gene expression.

Taken together, my thesis work has set up a stage for studying transcriptional regulation of *mks-1*.

References

Alexiev BA, Lin X, Sun CC, Brenner DS. 2006. Meckel-Gruber syndrome: pathologic manifestations, minimal diagnostic criteria, and differential diagnosis. *Archives of pathology & laboratory medicine* **130**(8): 1236-1238.

Altun ZF, Hall DH. 2012. Handbook of *C. elegans* Anatomy. WormAtlas.

Andersen EC, Horvitz HR. 2007. Two *C. elegans* histone methyltransferases repress lin-3 EGF transcription to inhibit vulval development. *Development* **134**(16): 2991-2999.

Baala L, Audollent S, Martinovic J, Ozilou C, Babron MC, Sivanandamoorthy S, Saunier S, Salomon R, Gonzales M, Rattenberry E et al. 2007. Pleiotropic effects of CEP290 (NPHP6) mutations extend to Meckel syndrome. *American journal of human genetics* **81**(1): 170-179.

Badano JL, Mitsuma N, Beales PL, Katsanis N. 2006. The ciliopathies: an emerging class of human genetic disorders. *Annual review of genomics and human genetics* **7**: 125-148.

Baker K, Beales PL. 2009. Making sense of cilia in disease: the human ciliopathies. *American journal of medical genetics Part C, Seminars in medical genetics* **151C**(4): 281-295.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell research* **21**(3): 381-395.

Bargmann CI. 2006. Chemosensation in *C. elegans*. *WormBook : the online review of C elegans biology*: 1-29.

Bargmann CI, Hartwieg E, Horvitz HR. 1993. Odorant-selective genes and neurons mediate olfaction in *C. elegans*. *Cell* **74**(3): 515-527.

Barr MM, Sternberg PW. 1999. A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*. *Nature* **401**(6751): 386-389.

Beckers A, Alten L, Viebahn C, Andre P, Gossler A. 2007. The mouse homeobox gene Noto regulates node morphogenesis, notochordal ciliogenesis, and left right patterning. *Proceedings of the National Academy of Sciences of the United States of America* **104**(40): 15765-15770.

Bentley DL. 2005. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Current opinion in cell biology* **17**(3): 251-256.

Bergmann C, Fliegauf M, Bruchle NO, Frank V, Olbrich H, Kirschner J, Schermer B, Schmedding I, Kispert A, Kranzlin B *et al.* 2008. Loss of nephrocystin-3 function can cause embryonic lethality, Meckel-Gruber-like syndrome, situs inversus, and renal-hepatic-pancreatic dysplasia. *American journal of human genetics* **82**(4): 959-970.

Bialas NJ, Inglis PN, Li C, Robinson JF, Parker JD, Healey MP, Davis EE, Inglis CD, Toivonen T, Cottell DC *et al.* 2009. Functional interactions between the ciliopathy-associated Meckel syndrome 1 (MKS1) protein and two novel MKS1-related (MKSR) proteins. *Journal of cell science* **122**(Pt 5): 611-624.

Blacque OE, Reardon MJ, Li C, McCarthy J, Mahjoub MR, Ansley SJ, Badano JL, Mah AK, Beales PL, Davidson WS *et al.* 2004. Loss of *C. elegans* BBS-7 and BBS-8 protein function results in cilia defects and compromised intraflagellar transport. *Genes & development* **18**(13): 1630-1642.

Bloodgood RA. 2009. From central to rudimentary to primary: the history of an underappreciated organelle whose time has come. The primary cilium. *Methods in cell biology* **94**: 3-52.

Boulin T, Etchberger JF, Hobert O. 2006. Reporter gene fusions. *WormBook : the online review of C elegans biology*: 1-23.

Brehm A, Miska EA, McCance DJ, Reid JL, Bannister AJ, Kouzarides T. 1998. Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**(6667): 597-601.

Brenner S. 1974. The genetics of Caenorhabditis elegans. *Genetics* **77**(1): 71-94.

Britton C, McKerrow JH, Johnstone IL. 1998. Regulation of the Caenorhabditis elegans gut cysteine protease gene cpr-1: requirement for GATA motifs. *Journal of molecular biology* **283**(1): 15-27.

Brody SL, Yan XH, Wuerffel MK, Song SK, Shapiro SD. 2000. Ciliogenesis and left-right axis defects in forkhead factor HFH-4-null mice. *American journal of respiratory cell and molecular biology* **23**(1): 45-51.

Brookes P, Lawley PD. 1961. The reaction of mono- and di-functional alkylating agents with nucleic acids. *The Biochemical journal* **80**(3): 496-503.

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. *Developmental biology* **368**(2): 415-426.

Cachero S, Simpson TI, Zur Lage PI, Ma L, Newton FG, Holohan EE, Armstrong JD, Jarman AP. 2011. The gene regulatory cascade linking proneural specification with differentiation in Drosophila sensory neurons. *PLoS biology* **9**(1): e1000568.

Chalfie M, Sulston J. 1981. Developmental genetics of the mechanosensory neurons of Caenorhabditis elegans. *Developmental biology* **82**(2): 358-370.

Chan HM, La Thangue NB. 2001. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *Journal of cell science* **114**(Pt 13): 2363-2373.

Chang S, Johnston RJ, Jr., Hobert O. 2003. A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes & development* **17**(17): 2123-2137.

Chen J, Knowles HJ, Hebert JL, Hackett BP. 1998. Mutation of the mouse hepatocyte nuclear factor/forkhead homologue 4 gene results in an absence of cilia and random left-right asymmetry. *The Journal of clinical investigation* **102**(6): 1077-1082.

Chen N, Mah A, Blacque OE, Chu J, Phgora K, Bakhoum MW, Newbury CR, Khattra J, Chan S, Go A *et al.* 2006. Identification of ciliary and ciliopathy genes in Caenorhabditis elegans through comparative genomics. *Genome biology* **7**(12): R126.

Chen N, Pai S, Zhao Z, Mah A, Newbury R, Johnsen RC, Altun Z, Moerman DG, Baillie DL, Stein LD. 2005. Identification of a nematode chemosensory gene family. *Proceedings of the National Academy of Sciences of the United States of America* **102**(1): 146-151.

Choksi SP, Lauter G, Swoboda P, Roy S. 2014. Switching on cilia: transcriptional networks regulating ciliogenesis. *Development* **141**(7): 1427-1441.

Chu JS, Baillie DL, Chen N. 2010. Convergent evolution of RFX transcription factors and ciliary genes predated the origin of metazoans. *BMC evolutionary biology* **10**: 130.

Chu JS, Tarailo-Graovac M, Zhang D, Wang J, Uyar B, Tu D, Trinh J, Baillie DL, Chen N. 2012. Fine tuning of RFX/DAF-19-regulated target gene expression through binding to multiple sites in Caenorhabditis elegans. *Nucleic acids research* **40**(1): 53-64.

Clevidence DE, Overdier DG, Tao W, Qian X, Pani L, Lai E, Costa RH. 1993. Identification of nine tissue-specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA-binding-domain family. *Proceedings of the National Academy of Sciences of the United States of America* **90**(9): 3948-3952.

Cole DG, Diener DR, Himelblau AL, Beech PL, Fuster JC, Rosenbaum JL. 1998. Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons. *The Journal of cell biology* **141**(4): 993-1008.

Consortium CeS. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**(5396): 2012-2018.

Consugar MB, Kubly VJ, Lager DJ, Hommerding CJ, Wong WC, Bakker E, Gattone VH, 2nd, Torres VE, Breuning MH, Harris PC. 2007. Molecular diagnostics of Meckel-Gruber syndrome highlights phenotypic differences between MKS1 and MKS3. *Human genetics* **121**(5): 591-599.

Coulondre C, Miller JH. 1977. Genetic studies of the lac repressor. III. Additional correlation of mutational sites with specific amino acid residues. *Journal of molecular biology* **117**(3): 525-567.

Cui C, Chatterjee B, Francis D, Yu Q, SanAgustin JT, Francis R, Tansey T, Henry C, Wang B, Lemley B *et al.* 2011. Disruption of Mks1 localization to the mother centriole causes cilia defects and developmental malformations in Meckel-Gruber syndrome. *Disease models & mechanisms* **4**(1): 43-56.

Culotti JG, Russell RL. 1978. Osmotic avoidance defective mutants of the nematode Caenorhabditis elegans. *Genetics* **90**(2): 243-256.

Dafinger C, Liebau MC, Elsayed SM, Hellenbroich Y, Boltshauser E, Korenke GC, Fabretti F, Janecke AR, Ebermann I, Nurnberg G *et al.* 2011. Mutations in KIF7 link Joubert syndrome with Sonic Hedgehog signaling and microtubule dynamics. *The Journal of clinical investigation* **121**(7): 2662-2667.

Dar PS, G.S.; Carter ,S.M.; Ferreira, J.C.; Nitowsky, H.M.; Gross, S.J. 2001. Prenatal diagnosis of Bardet-Biedl syndrome by targeted second-trimester sonography. *Ultrasound Obstet Gynecol* **17**: 354–356.

Davis EE, Katsanis N. 2012. The ciliopathies: a transitional model into systems biology of human genetic disease. *Current opinion in genetics & development* **22**(3): 290-303.

Davis EE, Zhang Q, Liu Q, Diplas BH, Davey LM, Hartley J, Stoetzel C, Szymanska K, Ramaswami G, Logan CV *et al.* 2011. TTC21B contributes both causal and modifying alleles across the ciliopathy spectrum. *Nature genetics* **43**(3): 189-196.

Dawe HR, Smith UM, Cullinane AR, Gerrelli D, Cox P, Badano JL, Blair-Reid S, Sriram N, Katsanis N, Attie-Bitach T *et al.* 2007. The Meckel-Gruber Syndrome proteins MKS1 and meckelin interact and are required for primary cilium formation. *Human molecular genetics* **16**(2): 173-186.

Delous M, Baala L, Salomon R, Laclef C, Vierkotten J, Tory K, Golzio C, Lacoste T, Besse L, Ozilou C *et al.* 2007. The ciliary gene RPGRIP1L is mutated in cerebello-oculo-renal syndrome (Joubert syndrome type B) and Meckel syndrome. *Nature genetics* **39**(7): 875-881.

den Hollander AI, Koenekoop RK, Yzer S, Lopez I, Arends ML, Voesenek KE, Zonneveld MN, Strom TM, Meitinger T, Brunner HG *et al.* 2006. Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis. *American journal of human genetics* **79**(3): 556-561.

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. *Respiratory research* **14**: 70.

Doitsidou M, Flames N, Lee AC, Boyanov A, Hobert O. 2008. Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. *Nature methods* **5**(10): 869-872.

Doitsidou M, Poole RJ, Sarin S, Bigelow H, Hobert O. 2010. *C. elegans* mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. *PloS one* **5**(11): e15435.

Dowdle WE, Robinson JF, Kneist A, Sirerol-Piquer MS, Frints SG, Corbit KC, Zaghloul NA, van Lijnschoten G, Mulders L, Verver DE *et al.* 2011. Disruption of a ciliary B9 protein complex causes Meckel syndrome. *American journal of human genetics* **89**(1): 94-110.

Durand B, Vandaele C, Spencer D, Pantalacci S, Couble P. 2000. Cloning and characterization of dRFX, the Drosophila member of the RFX family of transcription factors. *Gene* **246**(1-2): 285-293.

Dusenbery DB, Sheridan RE, Russell RL. 1975. Chemotaxis-defective mutants of the nematode Caenorhabditis elegans. *Genetics* **80**(2): 297-309.

Efimenko E, Bubb K, Mak HY, Holzman T, Leroux MR, Ruvkun G, Thomas JH, Swoboda P. 2005. Analysis of xbx genes in *C. elegans*. *Development* **132**(8): 1923-1934.

Emery P, Durand B, Mach B, Reith W. 1996. RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic acids research* **24**(5): 803-807.

Engels BM, Hutvagner G. 2006. Principles and effects of microRNA-mediated post-transcriptional gene regulation. *Oncogene* **25**(46): 6163-6169.

Evans JE, Snow JJ, Gunnarson AL, Ou G, Stahlberg H, McDonald KL, Scholey JM. 2006. Functional modulation of IFT kinesins extends the sensory repertoire of ciliated neurons in Caenorhabditis elegans. *The Journal of cell biology* **172**(5): 663-669.

Ezratty EJ, Stokes N, Chai S, Shah AS, Williams SE, Fuchs E. 2011. A role for the primary cilium in Notch signaling and epidermal differentiation during skin development. *Cell* **145**(7): 1129-1141.

Fernandez AG, Mis EK, Bargmann BO, Birnbaum KD, Piano F. 2010. Automated sorting of live *C. elegans* using laFACS. *Nature methods* **7**(6): 417-418.

Fisch C, Dupuis-Williams P. 2011. Ultrastructure of cilia and flagella - back to the future! *Biology of the cell / under the auspices of the European Cell Biology Organization* **103**(6): 249-270.

Forsythe E, Beales PL. 2013. Bardet-Biedl syndrome. *European journal of human genetics : EJHG* **21**(1): 8-13.

Freund CL, Gregory-Evans CY, Furukawa T, Papaioannou M, Looser J, Ploder L, Bellingham J, Ng D, Herbrick JA, Duncan A *et al.* 1997. Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) essential for maintenance of the photoreceptor. *Cell* **91**(4): 543-553.

Furukawa T, Morrow EM, Cepko CL. 1997. Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell* **91**(4): 531-541.

Gajiwala KS, Chen H, Cornille F, Roques BP, Reith W, Mach B, Burley SK. 2000. Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature* **403**(6772): 916-921.

Gazioglu N, Vural M, Seckin MS, Tuysuz B, Akpir E, Kuday C, Ilikkan B, Erginel A, Cenani A. 1998. Meckel-Gruber syndrome. *Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery* **14**(3): 142-145.

Gengyo-Ando K, Mitani S. 2000. Characterization of mutations induced by ethyl methanesulfonate, UV, and trimethylpsoralen in the nematode Caenorhabditis elegans. *Biochemical and biophysical research communications* **269**(1): 64-69.

Gerstein MB Lu ZJ Van Nostrand EL Cheng C Arshinoff BI Liu T Yip KY Robilotto R Rechtsteiner A Ikegami K *et al.* 2010. Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. *Science* **330**(6012): 1775-1787.

Gibbons IR. 1981. Cilia and flagella of eukaryotes. *The Journal of cell biology* **91**(3 Pt 2): 107s-124s.

Gilula NB, Satir P. 1972. The ciliary necklace. A ciliary membrane specialization. *The Journal of cell biology* **53**(2): 494-509.

Goetz SC, Anderson KV. 2010. The primary cilium: a signalling centre during vertebrate development. *Nature reviews Genetics* **11**(5): 331-344.

Grishkevich V, Hashimshony T, Yanai I. 2011. Core promoter T-blocks correlate with gene expression levels in *C. elegans*. *Genome research* **21**(5): 707-717.

Hall DH, Russell RL. 1991. The posterior nervous system of the nematode Caenorhabditis elegans: serial reconstruction of identified neurons and complete pattern of synaptic interactions. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **11**(1): 1-22.

Hedgecock EM, Culotti JG, Thomson JN, Perkins LA. 1985. Axonal guidance mutants of Caenorhabditis elegans identified by filling sensory neurons with fluorescein dyes. *Developmental biology* **111**(1): 158-170.

Helou J, Otto EA, Attanasio M, Allen SJ, Parisi MA, Glass I, Utsch B, Hashmi S, Fazzi E, Omran H *et al.* 2007. Mutation analysis of NPHP6/CEP290 in patients with Joubert syndrome and Senior-Loken syndrome. *Journal of medical genetics* **44**(10): 657-663.

Herman RK. 1984. Analysis of genetic mosaics of the nematode Caneorhabditis elegans. *Genetics* **108**(1): 165-180.

Hildebrandt F, Benzing T, Katsanis N. 2011. Ciliopathies. *The New England journal of medicine* **364**(16): 1533-1543.

Hillier LW, Coulson A, Murray JI, Bao Z, Sulston JE, Waterston RH. 2005. Genomics in *C. elegans*: so many genes, such a little worm. *Genome research* **15**(12): 1651-1660.

Hook EB. 1983. Down syndrome rates and relaxed selection at older maternal ages. *American journal of human genetics* **35**(6): 1307-1313.

Hopp K, Heyer CM, Hommerding CJ, Henke SA, Sundsbak JL, Patel S, Patel P, Consugar MB, Czarnecki PG, Gliem TJ *et al.* 2011. B9D1 is revealed as a novel Meckel syndrome (MKS) gene by targeted exon-enriched next-generation sequencing and deletion analysis. *Human molecular genetics* **20**(13): 2524-2534.

Huang M, Zhou Z, Elledge SJ. 1998. The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. *Cell* **94**(5): 595-605.

Hunt-Newbury R, Viveiros R, Johnsen R, Mah A, Anastas D, Fang L, Halfnight E, Lee D, Lin J, Lorch A *et al.* 2007. High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. *PLoS biology* **5**(9): e237.

Inglis PN, Ou G, Leroux MR, Scholey JM. 2007. The sensory cilia of Caenorhabditis elegans. *WormBook : the online review of C elegans biology*: 1-22.

Jorgensen EM, Mango SE. 2002. The art and design of genetic screens: caenorhabditis elegans. *Nature reviews Genetics* **3**(5): 356-369.

Kaelin WG, Jr. 2003. The von Hippel-Lindau gene, kidney cancer, and oxygen sensing. *Journal of the American Society of Nephrology : JASN* **14**(11): 2703-2711.

Karegowda LH, Shenoy PM, Sripathi S, Varman M. 2014. Joubert syndrome. *BMJ case reports* **2014**.

Karmous-Benailly H, Martinovic J, Gubler MC, Sirot Y, Clech L, Ozilou C, Auge J, Brahimi N, Etchevers H, Detrait E *et al.* 2005. Antenatal presentation of Bardet-Biedl syndrome may mimic Meckel syndrome. *American journal of human genetics* **76**(3): 493-504.

Kato M, de Lencastre A, Pincus Z, Slack FJ. 2009. Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during Caenorhabditis elegans development. *Genome biology* **10**(5): R54.

Kelly WG, Fire A. 1998. Chromatin silencing and the maintenance of a functional germline in Caenorhabditis elegans. *Development* **125**(13): 2451-2456.

Kiefer JC, Smith PA, Mango SE. 2007. PHA-4/FoxA cooperates with TAM-1/TRIM to regulate cell fate restriction in the *C. elegans* foregut. *Developmental biology* **303**(2): 611-624.

Kim SK, Shindo A, Park TJ, Oh EC, Ghosh S, Gray RS, Lewis RA, Johnson CA, Attie-Bittach T, Katsanis N *et al.* 2010. Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. *Science* **329**(5997): 1337-1340.

Kozminski KG, Johnson KA, Forscher P, Rosenbaum JL. 1993. A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proceedings of the National Academy of Sciences of the United States of America* **90**(12): 5519-5523.

Kroeger P, Stewart C, Schaap T, van Wijnen A, Hirshman J, Helms S, Stein G, Stein J. 1987. Proximal and distal regulatory elements that influence in vivo expression of a cell cycle-dependent human H4 histone gene. *Proceedings of the National Academy of Sciences of the United States of America* **84**(12): 3982-3986.

Kyttala M, Tallila J, Salonen R, Kopra O, Kohlschmidt N, Paavola-Sakki P, Peltonen L, Kestila M. 2006. MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome. *Nature genetics* **38**(2): 155-157.

Lee RC, Feinbaum RL, Ambros V. 1993. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**(5): 843-854.

Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN. 2004. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal* **23**(20): 4051-4060.

Leeuwenhoek A. 1677. Concerning little animals observed in rain-, well-, sea- and snow-water; as also in water wherein pepper had lain infused. *Philos Trans Lond* **12**: 821–831.

Leitch CC, Zaghloul NA, Davis EE, Stoetzel C, Diaz-Font A, Rix S, Alfadhel M, Lewis RA, Eyaid W, Banin E *et al.* 2008. Hypomorphic mutations in syndromic encephalocele genes are associated with Bardet-Biedl syndrome. *Nature genetics* **40**(4): 443-448.

Lemon B, Tjian R. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes & development* **14**(20): 2551-2569.

Levine M, Tjian R. 2003. Transcription regulation and animal diversity. *Nature* **424**(6945): 147-151.

Lichtsteiner S, Tjian R. 1993. Cloning and properties of the Caenorhabditis elegans TATA-box-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **90**(20): 9673-9677.

Mak HY, Nelson LS, Basson M, Johnson CD, Ruvkun G. 2006. Polygenic control of Caenorhabditis elegans fat storage. *Nature genetics* **38**(3): 363-368.

Malone EA, Thomas JH. 1994. A screen for nonconditional dauer-constitutive mutations in Caenorhabditis elegans. *Genetics* **136**(3): 879-886.

Mango SE, Lambie EJ, Kimble J. 1994. The pha-4 gene is required to generate the pharyngeal primordium of Caenorhabditis elegans. *Development* **120**(10): 3019-3031.

Maria BL, Boltshauser E, Palmer SC, Tran TX. 1999. Clinical features and revised diagnostic criteria in Joubert syndrome. *Journal of child neurology* **14**(9): 583-590; discussion 590-581.

Matsui T, Segall J, Weil PA, Roeder RG. 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *The Journal of biological chemistry* **255**(24): 11992-11996.

Mazet F, Yu JK, Liberles DA, Holland LZ, Shimeld SM. 2003. Phylogenetic relationships of the Fox (Forkhead) gene family in the Bilateria. *Gene* **316**: 79-89.

McKay SJ, Johnsen R, Khattra J, Asano J, Baillie DL, Chan S, Dube N, Fang L, Goszczynski B, Ha E *et al.* 2003. Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. *Cold Spring Harbor symposia on quantitative biology* **68**: 159-169.

Meckel JF. 1822. Beschreibung zweier, durch sehr ähnliche bildungsabweichungen entstelter geschwister. *Dtsch Arch Physiol* **7**: 99-172.

Mello C, Fire A. 1995. DNA transformation. *Methods in cell biology* **48**: 451-482.

Mello CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *The EMBO journal* **10**(12): 3959-3970.

Mencia A, Modamio-Hoybjor S, Redshaw N, Morin M, Mayo-Merino F, Olavarrieta L, Aguirre LA, del Castillo I, Steel KP, Dalmay T *et al.* 2009. Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. *Nature genetics* **41**(5): 609-613.

Mittermayer C, Lee A, Brugger PC. 2004. Prenatal diagnosis of the Meckel-Gruber syndrome from 11th to 20th gestational week. *Ultraschall in der Medizin* **25**(4): 275-279.

Mukhopadhyay A, Deplancke B, Walhout AJ, Tissenbaum HA. 2005. *C. elegans* tubby regulates life span and fat storage by two independent mechanisms. *Cell metabolism* **2**(1): 35-42.

Mukhopadhyay A, Pan X, Lambright DG, Tissenbaum HA. 2007a. An endocytic pathway as a target of tubby for regulation of fat storage. *EMBO reports* **8**(10): 931-938.

Mukhopadhyay S, Lu Y, Qin H, Lanjuin A, Shaham S, Sengupta P. 2007b. Distinct IFT mechanisms contribute to the generation of ciliary structural diversity in *C. elegans*. *The EMBO journal* **26**(12): 2966-2980.

Muller OF. 1786. Animalcula infusoria; fluvia tilia et marina, que detexit, systematice descripsit et ad

vivum delineari curavit. *Molleri, Havniae*.

Nalefski EA, Falke JJ. 1996. The C2 domain calcium-binding motif: structural and functional diversity. *Protein science : a publication of the Protein Society* **5**(12): 2375-2390.

Newton FG, zur Lage PI, Karak S, Moore DJ, Gopfert MC, Jarman AP. 2012. Forkhead transcription factor Fd3F cooperates with Rfx to regulate a gene expression program for mechanosensory cilia specialization. *Developmental cell* **22**(6): 1221-1233.

Niu W, Lu ZJ, Zhong M, Sarov M, Murray JI, Brdlik CM, Janette J, Chen C, Alves P, Preston E *et al.* 2011. Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. *Genome research* **21**(2): 245-254.

Oh SW, Mukhopadhyay A, Dixit BL, Raha T, Green MR, Tissenbaum HA. 2006. Identification of direct DAF-16 targets controlling longevity, metabolism and diapause by chromatin immunoprecipitation. *Nature genetics* **38**(2): 251-257.

Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A. 1993. Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. *Genetics* **135**(2): 385-404.

Okkema PG, Krause M. 2005. Transcriptional regulation. *WormBook : the online review of C elegans biology*: 1-40.

Orozco JT, Wedaman KP, Signor D, Brown H, Rose L, Scholey JM. 1999. Movement of motor and cargo along cilia. *Nature* **398**(6729): 674.

Orphanides G, Lagrange T, Reinberg D. 1996. The general transcription factors of RNA polymerase II. *Genes & development* **10**(21): 2657-2683.

Orphanides G, Reinberg D. 2002. A unified theory of gene expression. *Cell* **108**(4): 439-451.

Ou G, Blacque OE, Snow JJ, Leroux MR, Scholey JM. 2005. Functional coordination of intraflagellar transport motors. *Nature* **436**(7050): 583-587.

Paavola P, Salonen R, Weissenbach J, Peltonen L. 1995. The locus for Meckel syndrome with multiple congenital anomalies maps to chromosome 17q21-q24. *Nature genetics* **11**(2): 213-215.

Pachi A, Giancotti A, Torcia F, de Prosperi V, Maggi E. 1989. Meckel-Gruber syndrome: ultrasonographic diagnosis at 13 weeks' gestational age in an at-risk case. *Prenatal diagnosis* **9**(3): 187-190.

Pazour GJ, Baker SA, Deane JA, Cole DG, Dickert BL, Rosenbaum JL, Witman GB, Besharse JC. 2002. The intraflagellar transport protein, IFT88, is essential for vertebrate photoreceptor assembly and maintenance. *The Journal of cell biology* **157**(1): 103-113.

Perens EA, Shaham S. 2005. *C. elegans* daf-6 encodes a patched-related protein required for lumen formation. *Developmental cell* **8**(6): 893-906.

Perkins LA, Hedgecock EM, Thomson JN, Culotti JG. 1986. Mutant sensory cilia in the nematode Caenorhabditis elegans. *Developmental biology* **117**(2): 456-487.

Pflieger S, Lefebvre V, Causse M. 2001. The candidate gene approach in plant genetics: a review. *Mol Breeding* **7**(4): 275-291.

Powell-Coffman JA, Knight J, Wood WB. 1996. Onset of *C. elegans* gastrulation is blocked by inhibition of embryonic transcription with an RNA polymerase antisense RNA. *Developmental biology* **178**(2): 472-483.

Praetorius HA, Spring KR. 2003. Removal of the MDCK cell primary cilium abolishes flow sensing. *The Journal of membrane biology* **191**(1): 69-76.

Pujol N, Torregrossa P, Ewbank JJ, Brunet JF. 2000. The homeodomain protein CePHOX2/CEH-17 controls antero-posterior axonal growth in *C. elegans*. *Development* **127**(15): 3361-3371.

Pulak R. 2006. Techniques for analysis, sorting, and dispensing of *C. elegans* on the COPAS flow-sorting system. *Methods in molecular biology* **351**: 275-286.

Putoux A, Thomas S, Coene KL, Davis EE, Alanay Y, Ogur G, Uz E, Buzas D, Gomes C, Patrier S *et al.* 2011. KIF7 mutations cause fetal hydrolethalus and acrocallosal syndromes. *Nature genetics* **43**(6): 601-606.

Reece-Hoyes JS, Deplancke B, Shingles J, Grove CA, Hope IA, Walhout AJ. 2005. A compendium of Caenorhabditis elegans regulatory transcription factors: a resource for mapping transcription regulatory networks. *Genome biology* **6**(13): R110.

Reece-Hoyes JS, Marian Walhout AJ. 2012. Yeast one-hybrid assays: a historical and technical perspective. *Methods* **57**(4): 441-447.

Reid LH, Gregg RG, Smithies O, Koller BH. 1990. Regulatory elements in the introns of the human HPRT gene are necessary for its expression in embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **87**(11): 4299-4303.

Reinke V, Krause M, Okkema P. 2013. Transcriptional regulation of gene expression in *C. elegans*. *WormBook : the online review of C elegans biology*: 1-34.

Reiter JF, Blacque OE, Leroux MR. 2012. The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization. *EMBO reports* **13**(7): 608-618.

Reith W, Herrero-Sanchez C, Kobr M, Silacci P, Berte C, Barras E, Fey S, Mach B. 1990. MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain. *Genes & development* **4**(9): 1528-1540.

Riazuddin SA, Iqbal M, Wang Y, Masuda T, Chen Y, Bowne S, Sullivan LS, Waseem NH, Bhattacharya S, Daiger SP *et al.* 2010. A splice-site mutation in a retina-specific exon of BBS8 causes nonsyndromic retinitis pigmentosa. *American journal of human genetics* **86**(5): 805-812.

Riddle DL, Swanson MM, Albert PS. 1981. Interacting genes in nematode dauer larva formation. *Nature* **290**(5808): 668-671.

Rosenbluth RE, Cuddeford C, Baillie DL. 1985. Mutagenesis in Caenorhabditis elegans. II. A spectrum of mutational events induced with 1500 r of gamma-radiation. *Genetics* **109**(3): 493-511.

Roume J, Genin E, Cormier-Daire V, Ma HW, Mehaye B, Attie T, Razavi-Encha F, Fallet-Bianco C, Buenerd A, Clerget-Darpoux F *et al.* 1998. A gene for Meckel syndrome maps to chromosome 11q13. *American journal of human genetics* **63**(4): 1095-1101.

Saito H, Maki RA, Clayton LK, Tonegawa S. 1983. Complete primary structures of the E beta chain and gene of the mouse major histocompatibility complex. *Proceedings of the National Academy of Sciences of the United States of America* **80**(18): 5520-5524.

Salonen R, Norio R. 1984. The Meckel syndrome in Finland: epidemiologic and genetic aspects. *American journal of medical genetics* **18**(4): 691-698.

Salonen R, Paavola P. 1998. Meckel syndrome. *Journal of medical genetics* **35**(6): 497-501.

Sang L, Miller JJ, Corbit KC, Giles RH, Brauer MJ, Otto EA, Baye LM, Wen X, Scales SJ, Kwong M *et al.* 2011. Mapping the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways. *Cell* **145**(4): 513-528.

Sarin S, Antonio C, Tursun B, Hobert O. 2009. The *C. elegans* Tailless/TLX transcription factor nhr-67 controls neuronal identity and left/right asymmetric fate diversification. *Development* **136**(17): 2933-2944.

Sarin S, Bertrand V, Bigelow H, Boyanov A, Doitsidou M, Poole RJ, Narula S, Hobert O. 2010. Analysis of multiple ethyl methanesulfonate-mutagenized Caenorhabditis elegans strains by whole-genome sequencing. *Genetics* **185**(2): 417-430.

Satir P, Christensen ST. 2008. Structure and function of mammalian cilia. *Histochemistry and cell biology* **129**(6): 687-693.

Schafer JC, Winkelbauer ME, Williams CL, Haycraft CJ, Desmond RA, Yoder BK. 2006. IFTA-2 is a conserved cilia protein involved in pathways regulating longevity and dauer formation in Caenorhabditis elegans. *Journal of cell science* **119**(Pt 19): 4088-4100.

Senti G, Swoboda P. 2008. Distinct isoforms of the RFX transcription factor DAF-19 regulate ciliogenesis and maintenance of synaptic activity. *Molecular biology of the cell* **19**(12): 5517-5528.

Sepulveda W, Sebire NJ, Souka A, Snijders RJ, Nicolaides KH. 1997. Diagnosis of the Meckel-Gruber syndrome at eleven to fourteen weeks' gestation. *American journal of obstetrics and gynecology* **176**(2): 316-319.

Seydoux G, Dunn MA. 1997. Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. *Development* **124**(11): 2191-2201.

Seydoux G, Fire A. 1994. Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans. *Development* **120**(10): 2823-2834.

Shaheen R, Faqeih E, Alshammari MJ, Swaid A, Al-Gazali L, Mardawi E, Ansari S, Sogaty S, Seidahmed MZ, AlMotairi MI *et al.* 2013. Genomic analysis of Meckel-Gruber syndrome in Arabs reveals marked genetic heterogeneity and novel candidate genes. *European journal of human genetics : EJHG* **21**(7): 762-768.

Sharpey W. 1835. Cilia. *Cyclopoedia of Anatomy and Physiology (R B Todd, ed)* **1**: 606–638.

Shi Y, Mello C. 1998. A CBP/p300 homolog specifies multiple differentiation pathways in Caenorhabditis elegans. *Genes & development* **12**(7): 943-955.

Singh N, Han M. 1995. sur-2, a novel gene, functions late in the let-60 ras-mediated signaling pathway during Caenorhabditis elegans vulval induction. *Genes & development* **9**(18): 2251-2265.

Smith UM, Consugar M, Tee LJ, McKee BM, Maina EN, Whelan S, Morgan NV, Goranson E, Gissen P, Lilliquist S *et al.* 2006. The transmembrane protein meckelin (MKS3) is mutated in Meckel-Gruber syndrome and the wpk rat. *Nature genetics* **38**(2): 191-196.

Snow JJ, Ou G, Gunnarson AL, Walker MR, Zhou HM, Brust-Mascher I, Scholey JM. 2004. Two anterograde intraflagellar transport motors cooperate to build sensory cilia on *C. elegans* neurons. *Nature cell biology* **6**(11): 1109-1113.

Sorokin S. 1962. Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells. *The Journal of cell biology* **15**: 363-377.

Sorokin SP. 1968. Reconstructions of centriole formation and ciliogenesis in mammalian lungs. *Journal of cell science* **3**(2): 207-230.

Sotelo JR, Trujillo-Cenoz O. 1958. Electron microscope study on the development of ciliary components of the neural epithelium of the chick embryo. *Zeitschrift fur Zellforschung und mikroskopische Anatomie* **49**(1): 1-12.

Starich TA, Herman RK, Kari CK, Yeh WH, Schackwitz WS, Schuyler MW, Collet J, Thomas JH, Riddle DL. 1995. Mutations affecting the chemosensory neurons of Caenorhabditis elegans. *Genetics* **139**(1): 171-188.

Stimpson KM, Sullivan BA. 2010. Epigenomics of centromere assembly and function. *Current opinion in cell biology* **22**(6): 772-780.

Stoeckius M, Maaskola J, Colombo T, Rahn HP, Friedlander MR, Li N, Chen W, Piano F, Rajewsky N. 2009. Large-scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression. *Nature methods* **6**(10): 745-751.

Stoetzel C, Laurier V, Davis EE, Muller J, Rix S, Badano JL, Leitch CC, Salem N, Chouery E, Corbani S *et al.* 2006. BBS10 encodes a vertebrate-specific chaperonin-like protein and is a major BBS locus. *Nature genetics* **38**(5): 521-524.

Stubbs JL, Oishi I, Izpisua Belmonte JC, Kintner C. 2008. The forkhead protein Foxj1 specifies node-like cilia in Xenopus and zebrafish embryos. *Nature genetics* **40**(12): 1454-1460.

Stubbs JL, Vladar EK, Axelrod JD, Kintner C. 2012. Multicilin promotes centriole assembly and ciliogenesis during multiciliate cell differentiation. *Nature cell biology* **14**(2): 140-147.

Sulston JE, Albertson DG, Thomson JN. 1980. The Caenorhabditis elegans male: postembryonic development of nongonadal structures. *Developmental biology* **78**(2): 542-576.

Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental biology* **56**(1): 110-156.

Swoboda P, Adler HT, Thomas JH. 2000. The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Molecular cell* **5**(3): 411-421.

Tabor HK, Risch NJ, Myers RM. 2002. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature reviews Genetics* **3**(5): 391-397.

Tallila J, Jakkula E, Peltonen L, Salonen R, Kestila M. 2008. Identification of CC2D2A as a Meckel syndrome gene adds an important piece to the ciliopathy puzzle. *American journal of human genetics* **82**(6): 1361-1367.

Tammachote R, Hommerding CJ, Sinders RM, Miller CA, Czarnecki PG, Leightner AC, Salisbury JL, Ward CJ, Torres VE, Gattone VH, 2nd *et al.* 2009. Ciliary and centrosomal defects associated with mutation and depletion of the Meckel syndrome genes MKS1 and MKS3. *Human molecular genetics* **18**(17): 3311-3323.

Thomas J, Morle L, Soulavie F, Laurencon A, Sagnol S, Durand B. 2010. Transcriptional control of genes involved in ciliogenesis: a first step in making cilia. *Biology of the cell / under the auspices of the European Cell Biology Organization* **102**(9): 499-513.

Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, Au V, Chaudhry I, Fernando L, Hutter H *et al.* 2013. The million mutation project: a new approach to genetics in Caenorhabditis elegans. *Genome research* **23**(10): 1749-1762.

Troemel ER, Sagasti A, Bargmann CI. 1999. Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell* **99**(4): 387-398.

Uchida O, Nakano H, Koga M, Ohshima Y. 2003. The *C. elegans* che-1 gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons. *Development* **130**(7): 1215-1224.

Wagmaister JA, Miley GR, Morris CA, Gleason JE, Miller LM, Kornfeld K, Eisenmann DM. 2006. Identification of cis-regulatory elements from the *C. elegans* Hox gene lin-39 required for embryonic expression and for regulation by the transcription factors LIN-1, LIN-31 and LIN-39. *Developmental biology* **297**(2): 550-565.

Warburton-Pitt SR, Jauregui AR, Li C, Wang J, Leroux MR, Barr MM. 2012. Ciliogenesis in Caenorhabditis elegans requires genetic interactions between ciliary middle segment localized NPHP-2 (inversin) and transition zone-associated proteins. *Journal of cell science* **125**(Pt 11): 2592-2603.

Ward S, Thomson N, White JG, Brenner S. 1975. Electron microscopical reconstruction of the anterior sensory anatomy of the nematode Caenorhabditis elegans.?2UU. *The Journal of comparative neurology* **160**(3): 313-337.

Ware RW, Clark, D., Crossland, K., and Russell, R.L. 1975. The nerve ring of the nematode Caenorhabditis elegans: sensory input and motor output. . *J Comp Neurol* **162**: 71–110.

Waters AM, Beales PL. 1993-2003 (updated 2011). Bardet–Biedl Syndrome. In *GeneReviews*, Internet.

Waters JC. 2009. Accuracy and precision in quantitative fluorescence microscopy. *The Journal of cell biology* **185**(7): 1135-1148.

Weatherbee SD, Niswander LA, Anderson KV. 2009. A mouse model for Meckel syndrome reveals Mks1 is required for ciliogenesis and Hedgehog signaling. *Human molecular genetics* **18**(23): 4565-4575.

Wheway G, Abdelhamed Z, Natarajan S, Toomes C, Inglehearn C, Johnson CA. 2013. Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel-Gruber syndrome. *Developmental biology* **377**(1): 55-66.

Williams CL, Li C, Kida K, Inglis PN, Mohan S, Semenec L, Bialas NJ, Stupay RM, Chen N, Blacque OE *et al.* 2011. MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis. *The Journal of cell biology* **192**(6): 1023-1041.

Williams CL, Winkelbauer ME, Schafer JC, Michaud EJ, Yoder BK. 2008. Functional redundancy of the B9 proteins and nephrocystins in Caenorhabditis elegans ciliogenesis. *Molecular biology of the cell* **19**(5): 2154-2168.

Wittenburg N, Baumeister R. 1999. Thermal avoidance in Caenorhabditis elegans: an approach to the study of nociception. *Proceedings of the National Academy of Sciences of the United States of America* **96**(18): 10477-10482.

Yu X, Ng CP, Habacher H, Roy S. 2008. Foxj1 transcription factors are master regulators of the motile ciliogenic program. *Nature genetics* **40**(12): 1445-1453.

Zhao Z, Fang L, Chen N, Johnsen RC, Stein L, Baillie DL. 2005. Distinct regulatory elements mediate similar expression patterns in the excretory cell of Caenorhabditis elegans. *The Journal of biological chemistry* **280**(46): 38787-38794.

Zhao Z, Thomas JH, Chen N, Sheps JA, Baillie DL. 2007. Comparative genomics and adaptive selection of the ATP-binding-cassette gene family in caenorhabditis species. *Genetics* **175**(3): 1407-1418.

Appendix.

Candidate gene approach

Introduction: Candidate genes

Candidate gene approach is widely used in genetics studies. In candidate gene approach, it is important to hypothesize possible candidate genes using statistical models. Then the candidates are identified with genetics methods. The aim is to find out the functionally important genes, which work in a signalling pathway, or regulate other genes. In studies of genetic disease, candidate gene approach is used to identify genes in complex genetic diseases (Tabor *et al.* 2002). In plant genetics, candidate gene approach is applied to identify qualitative trait locus (QTL) (Pflieger *et al.* 2001).

In my study, I used this method, in addition to genetic screening, for two reasons: First, I had a list of transcriptional factors that affect gene expression in ciliated neurons or other neurons. Mutants of these transcriptional factors are available in the CGC or NBRP. I would know whether one of these candidates regulates *mks-1* by crossing *mks-1* promoter driven GFP reporter into the mutant background. Second, I was learning how to carry out genetic cross. It was a good practice to work on multiple genetic crosses by this method. So this piece of work was done at the beginning of my master study.

Dr. Jack Chen searched the SAGE profiling database. SAGE is Serial Analysis of Gene Expression. It is used for quantitative and qualitative analysis of expressed RNA. *C. elegans* SAGE database is available to run at http://elegans.bcgsc.ca/home/sage.html (McKay *et al.* 2003). He got a list of 6 genes that code for known transcriptional factors in *C. elegans* (Table A1). He ordered mutant strains of all 6 genes. I crossed the *mks-1* promoter driven GFP reporter into these 6 mutants' background.

Table A 1. Information of six candidate genes

Gene name	Location	Functional domain	Expression Profiles	Cellular function
tag-97	Chr X	ETS protein domains	ALM neurons CAN neurons Amphids and phasmids neurons Excretory cells	NA
nhr-20	Chr III	ligand-binding, core Zinc finger, nuclear hormone receptor type	AFD neurons Pan-neural cells Pharyngeal marginal cells	Nuclear Hormone Receptor
ceh-31	Chr X	homeobox domain	Amphids neurons	NA
F22D6.2	Chr I	Zinc finger, A20 type Zinc finger, AN1 type	AFD neurons ASER neurons Ciliated neurons Pan-neural cells Pharyngeal marginal cells Pharynx cells	NA
zip-5	Chr V	Basic-leucine zipper domain	Head and tail neurons Head muscles and intestines	bZip transcription factor
nhr-4	Chr IV	ligand-binding, core Steroid hormone receptor Zinc finger, nuclear hormone receptor type	AFD neurons ASER neurons Ciliated neurons Pan-neural cells Pharyngeal marginal cells Pharynx cells	Nuclear hormone receptor

- The information is collected from SAGE profiling database (http://elegans.bcgsc.ca/home/sage.html) and Wormbase (http:// www.wormbase.org) . SAGE data were obtained from the Genome BC *C. elegans* Gene Expression Consortium http://elegans.bcgsc.bc.ca/. These SAGE data were produced at the Michael Smith Genome Sciences Centre with funding from Genome Canada

Materials and Methods

Table A2 shows the information of 6 mutant strains that Dr. Jack Chen ordered (Table A2). Before I introduced *mks-1* reporter to mutant, each mutant strain was crossed with N2 males for four times to clean the background mutations (Figure A1). The 4x cross (with N2) was done in cooperation with Zhaozhao Qin. I backcrossed *ok286* mutant and *tm1059* mutant. Zhaozhao Qin backcrossed *tm1068*, *gk646, ok1685* and *tm239* mutants (Table A2). All six mutants carry large deletion. To detect mutant alleles in cross

progeny, primers were designed to genotype the mutant alleles. There were three working primers: a forward and a reverse that work outside the deletion region, and a deletion forward that reside inside the deletions. In wildtype, two amplicons were generated. In mutant, one amplicon was generated. Primer binding positions in the genome for detection of mutant allele ok286 (tag-97) are shown as an example (Figure A2). The sequence for all six groups of primers and expected PCR products are: 1. *tag-97 (ok286)*: genoF: 5' ACACATGTGGCAAAACCAGA 3'; genodelF: 5' TTGTCGCCAATGCTCAAATA 3'; genoR 5' TTTTCGGGGAAAACTGAATG 3'; N2 has two bands of 2942bp and 916bp; *ok286* mutant has one band of 582bp; heterozygous worm has all three above bands. 2. *nhr-20 (tm1059)*: genoF: 5' GGCTGTGTACATTTCGCTCA 3'; genodelF: 5' TTATGCAGCGCTCAACATTT 3'; genoR 5' GGGTCTTGCCACCACTAAAA 3'; N2 has two bands of 1145bp and 807bp; *tm1059* mutant has one band of 594bp; heterozygous worm has all three above bands. 3. *ceh-31 (tm0239)*: genoF: 5' TTCCAACCCACTCATTCACA 3'; genodelF: 5' AACGCTCCAGTAGTCACATGG 3'; genoR 5' GCTTGTCGTTTCCATTTGGT 3'; N2 has two bands of 1213bp and 802bp; *tm0239* mutant has one band of 500bp; heterozygous worm has all three above bands. 4. *F22D6.2 (ok1685)*: genoF: 5' CTTGTGACCTGCCTCTCCTC 3'; genodelF: 5' GCTGCTGTCGTTTCGTTTCT 3'; genoR 5' TCAACTCACGTCACCGAATC 3'; N2 has two bands of 1473bp and 871bp; *ok1685* mutant has one band of 432bp; heterozygous worm has all three above bands. 5. *zip-5 (gk646)*: genoF: 5' TGGTTTGGTCCCACTTTTTC 3'; genodelF: 5' CCCACCACATCATGATCACTT 3'; genoR 5' GGGTTTATTTCCACCCCAAT3'; N2 has two bands of 2300bp and 541bp; *gk646* mutant has one band of 293bp; heterozygous worm has all three above bands. 6. *nhr-4 (tm1068)*: genoF: 5' GATCCGAGATTTGACGCATT 3'; genodelF: 5' TCAGCTGACACTTCTCCATCA 3'; genoR 5' AATAATCCTGGAGGGGTTGG 3'; N2 has two bands of 1083bp and 813bp; *tm1068* mutant has one band of 519bp; heterozygous worm has all three above bands. Zhaozhao Qin kindly designed the primers.

Table A 2. Information of six mutant strains

Gene name	Allele name	Mutant strain name	4X backcrossed clean strain name
tag-97	*ok286*	RB568	JNC215
nhr-20	*tm1059*	FX01059	JNC210
ceh-31	*tm0239*	FX00239	JC218
F22D6.2	*ok1685*	VC1240	JNC223
zip-5	*gk646*	VC1392	JNC204
nhr-4	*tm1068*	FX01068	JNC211

- The mutant strains were ordered from CGC and NBRP. After backcross, each strain was assigned with JNC number. The backcross was done in cooperation with Zhaozhao Qin.

Then I crossed the *mks-1* promoter driven GFP transgene with the cleaned mutants. I isolated homozygous reporter in homozygous mutant alleles background (Figure A3). The heterozygous F1 progeny with GFP signal was picked up. In F2 progeny, homozygous mutant worms were detected with PCR genotyping, using primers mentioned before. Theoretically a quarter of these mutant homozygous worms would be GFP homozygous. The GFP reporter was detected with signal observation under

dissecting microscope. If the transcriptional factor did not affect *mks-1* expression, the GFP homozygous worms were detected by GFP signals. F3 generations were observed, and I picked up plates that all worms have GFP. If the reporter was silenced because of the mutant alleles, all worms did not have GFP signal. In this case, PCR genotyping of GFP was used to detect GFP homozygous line (see Chapter 2 methods). After detecting GFP and mutant allele homozygous worms, I assigned each line a JNC number. Then I observed the *mks-1* promoter driven GFP expression in mutants under confocal microscope.

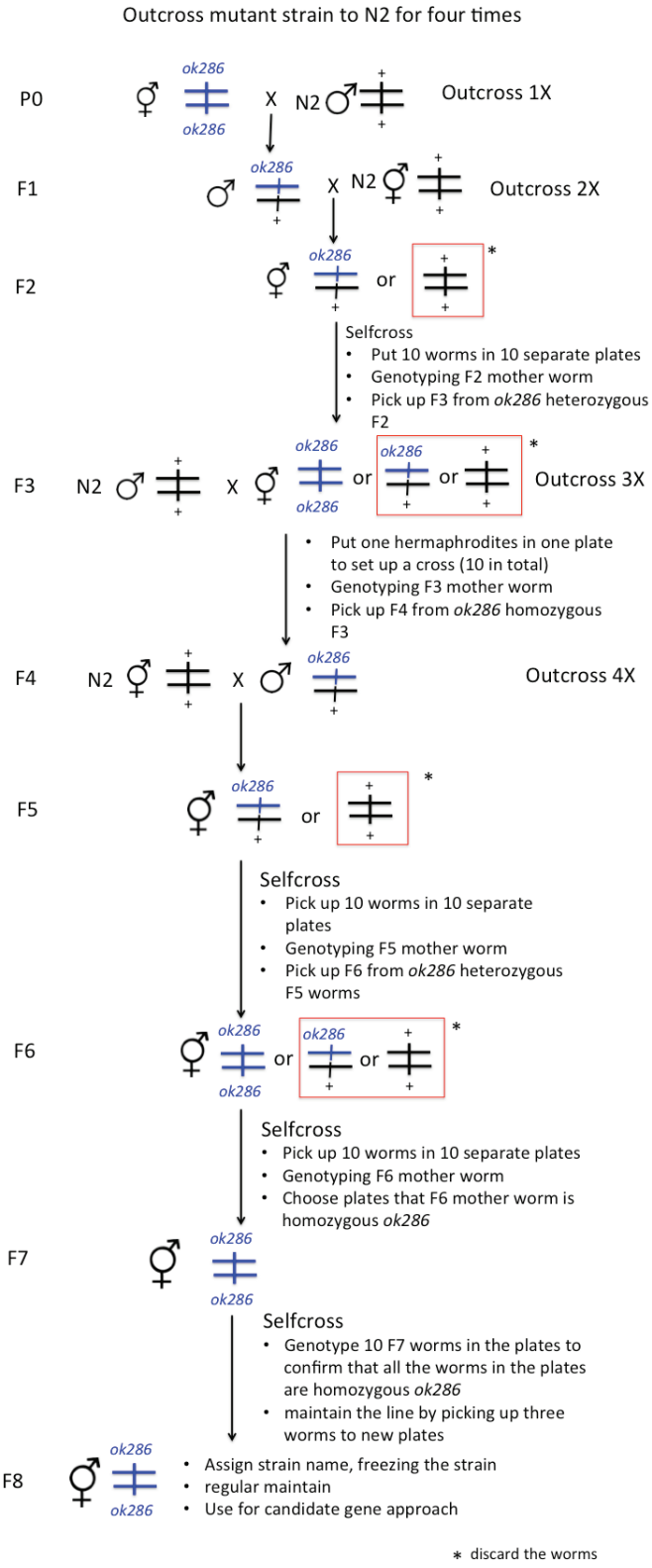
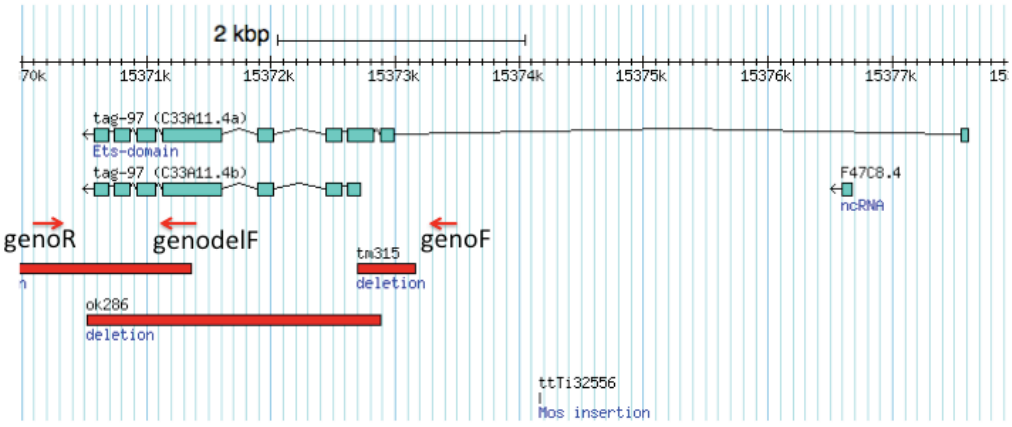


Figure A 1. Strategy for backcross mutants to N2 for four times

Note. Use *tag-97 (ok286)* mutant (strain RB568) as an example.

A



PCR Expected results with 1.5 min elongation time:
Wildtype (N2): around ~900bp band
Mutant (RB568): around ~550bp band

B

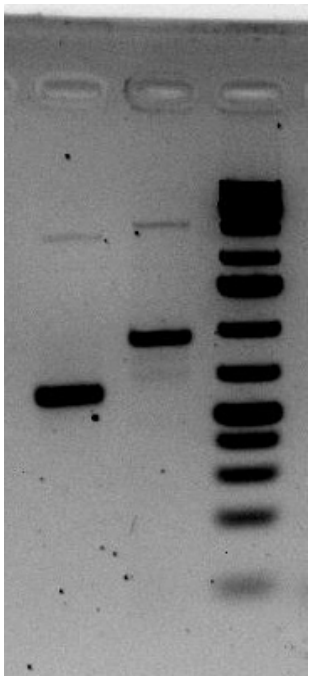


Figure A 2. Primers to detect mutant alleles *tag-97 (ok286)*

Note. A: Primers labeled in Wormbase Gbrowse view of Gene model showing *tag-97 (ok286).* Three primers were used, and one resided inside the *ok286* deletion region. PCR results to genotype *tag-97 (ok286).* Gene model image is downloaded from wormbase genome browser (http:://www.wormbase.org). B: First lane: ~550 bp band in *ok286* mutant; Second lane: ~900 bp band in N2; Third lane: GeneRuler 1 kb plus DNA ladder.

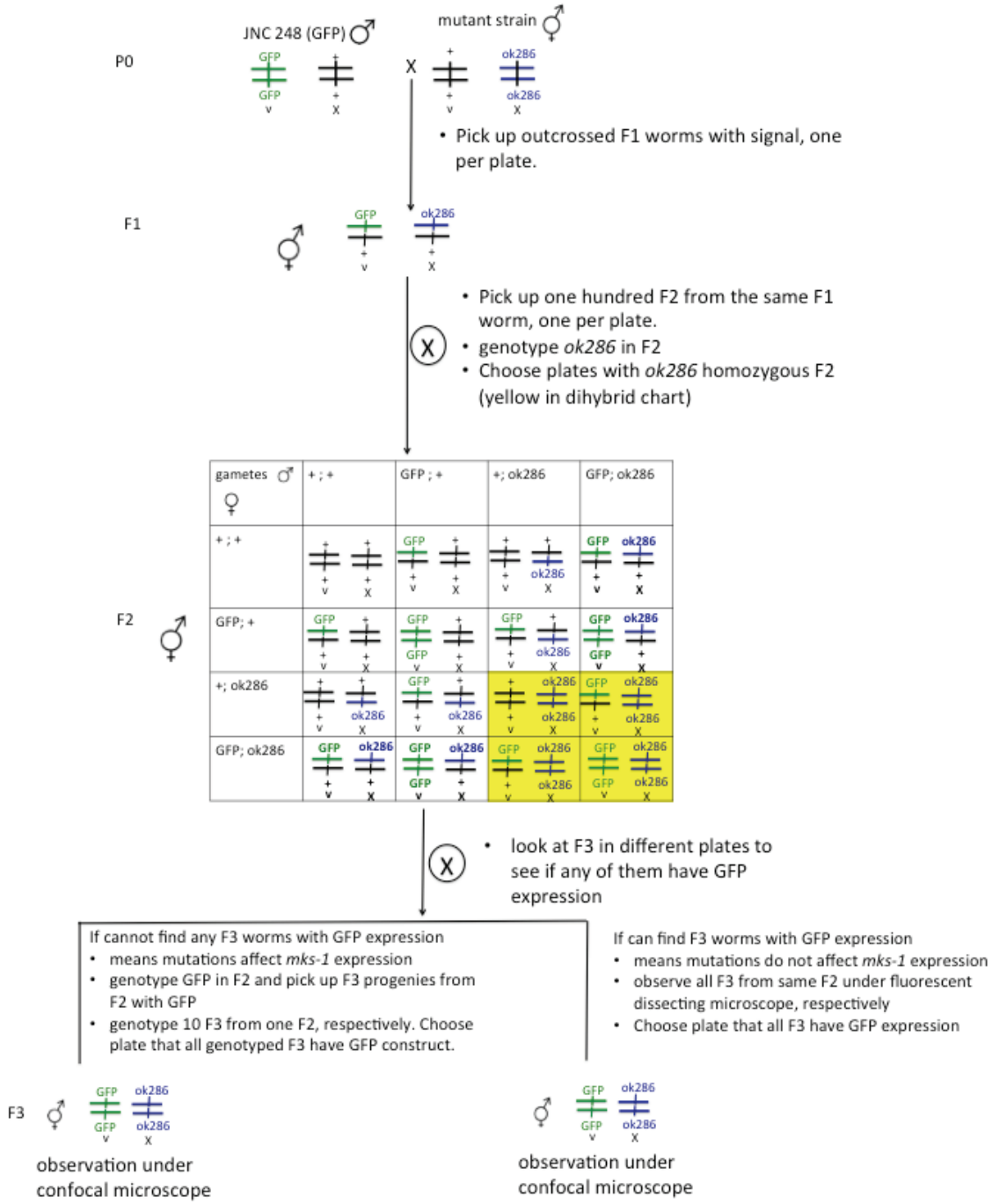


Figure A 3. Strategy to cross *mks-1* promoter driven GFP reporter to mutant background

Note. *ok286* mutant is used as an example. *ok286* is large deletions in *tag-97* gene.

Results and Discussions

I generated 6 strains of *mks-1* promoter driven GFP reporter in the mutant of *tag-97 (ok286), nhr-20 (tm1059), nhr-4 (tm1068), zip-5 (gk646), F22D6.2 (ok1658) and ceh-31 (tm0239)* respectively (Table A3). In all mutant backgrounds, *mks-1* promoter driven GFP expression was unchanged, compared to GFP expression in wildtype (Figure A4). None of the six candidate genes affect the expression of *mks-1*. Thus, t*ag-97, nhr-20, nhr-4, zip-5, F22D6.2* and *ceh-31* gene are regulating *mks-1*.

Table A 3. Infromation of the six strains with *mks-1* promoter driven GFP reporter in the mutant background

Strain name	Allele	Genotype	Mutated gene	GFP expression*
JNC275	*dotIs248; ok286*	*dotIs248[pmks-1::gfp+dpy-5(+)] tag-97(ok286) X*	*tag-97*	Normal
JNC257	*dotIs248; tm1059*	*dotIs248[pmks-1::gfp+dpy-5(+)] nhr-20(tm1059) III*	*nhr-20*	Normal
JNC276	*dotIs248; tm0239*	*dotIs248[pmks-1::gfp+dpy-5(+)] ceh-31(tm0239) X*	*ceh-31*	Normal
JNC277	*dotIs248; ok1685*	*dotIs248[pmks-1::gfp+dpy-5(+)] F22D6.2(ok1685) I*	*F22D6.2*	Normal
JNC281	*dotIs248; gk646*	*dotIs248[pmks-1::gfp+dpy-5(+)] zip-5(gk646) V*	*zip-5*	Normal
JNC280	*dotIs248; tm1068*	*dotIs248[pmks-1::gfp+dpy-5(+)] nhr-4(tm1068) IV*	*nhr-4*	Normal

* GFP expression refers to expression strength compared with GFP expression in JNC248 reporter strain. Normal means similar expression. Weak means the expression is weaker than JNC248 expression. Strong means the expression is stronger than JNC248 expression. None means could not see any expression under dissecting microscope.

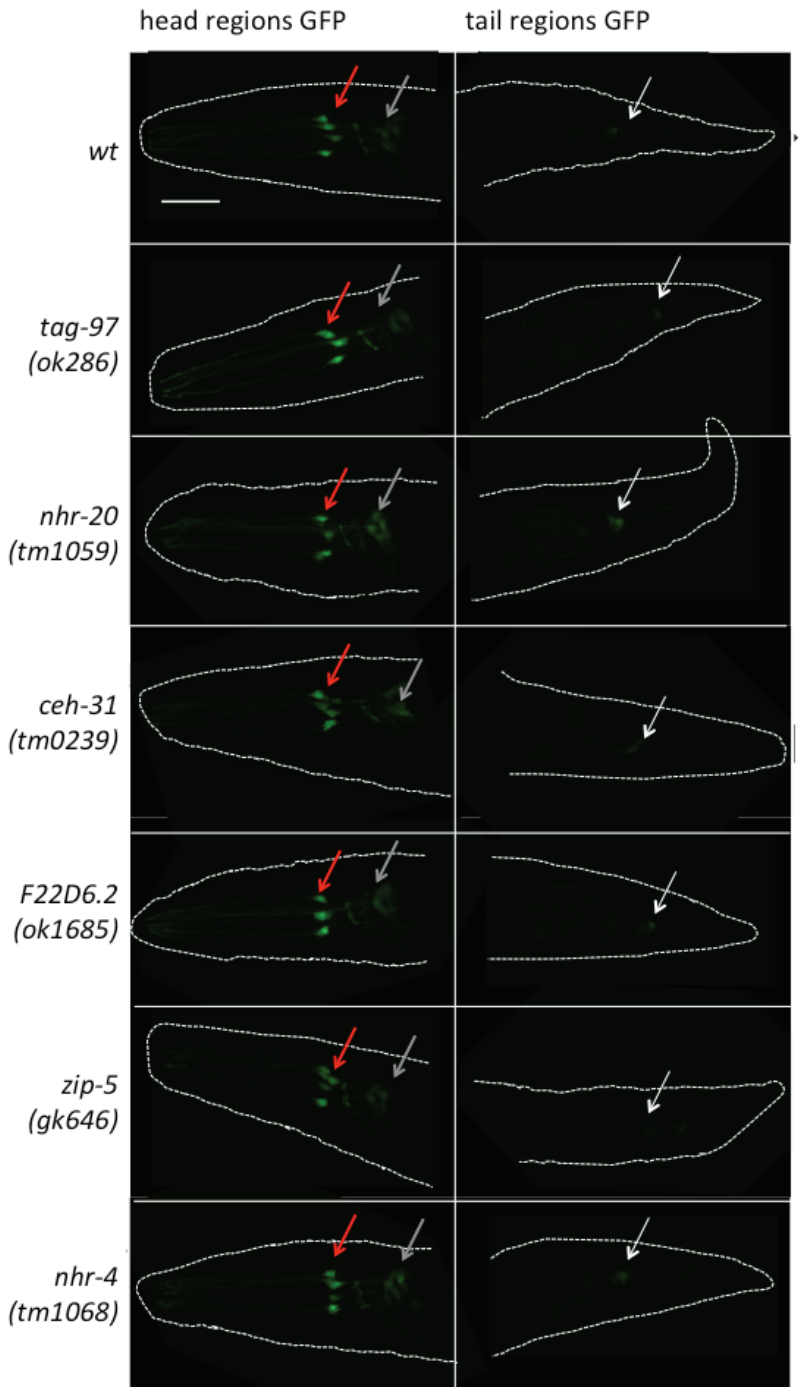


Figure A 4. Comparison for *mks-1* promoter driven GFP expression in mutants and in wildtype transgene

Note. A, B: *mks-1* promoter driven GFP expression in the wildtype transgenic worm (JNC248 strain). C-N: *mks-1* promoter driven GFP expression in different mutants. Read arrowheads: inner labial neurons; grey arrowheads: amphids; white arrowhead: phasmids. All images were taken under same exposure time 354 ms and sensitivity. No contrast enhancement was applied to images. Scale bar: 25 µm

Conclusion

Although I did not find the transcription factors of *mks-1* with candidate gene approach, the results are significant to report. When people work on transcriptional regulation of *mks-1* in *C. elegans* in the future, they can exclude these six genes: t*ag-97, nhr-20, nhr-4, zip-5, F22D6.2* and *ceh-31*.

Moreover, I tested candidate gene approach protocol that can be applied to genetic identification of candidate genes from genetic screens. After identifying mutations in EMS mutants, Chen lab colleagues will use candidate gene approach to genetically identify transcription factors. I proposed that it is applicable to combine methods of genetic screening and candidate gene approach. In the future, researchers will generate candidate genes' list from genetic screening and bioinformatics analysis. Then they will identify targeted genes by candidate gene approach.