From Department of Biosciences and Nutrition, Karolinska Institutet,
and School of Life Sciences, Södertörns Högskola
Stockholm, Sweden

Regulatory function of homeobox genes in the development of *Caenorhabditis elegans*

Yong Guang Tong, M.D.

Karolinska Institutet

Stockholm 2010

Published by Karolinska Institutet. Printed by Larserics Print AB.

ISBN 978-91-7457-018-2

To my parents and Cindy

献给我的父母和女儿

ABSTRACT

The nematode worm *Caenorhabditis elegans* has been widely used as a genetic model for over 40 years to investigate developmental control genes. In this thesis, I studied the roles of several homeobox genes and a novel RNA binding protein (RBD) in the development of *C. elegans* to understand the function of these genes in higher organisms. Homeobox genes are transcriptional regulators that are highly conserved in evolution and play important roles in the development and differentiation of multicellular organisms. The RBD is one of the most common protein domains in eukaryotes, and genes encoding this domain play roles in a wide variety of post-transcriptional gene regulation processes.

In paper I, we characterized a novel protein, RNA binding domain-1 (RBD-1), which is involved in ribosome biogenesis. This protein contains six consensus RNA-binding domains and is conserved as to sequence, domain organization, and subcellular localization from yeast to human. RBD-1 is essential for the development of *C. elegans*. The RNAi experiments using the cDNA of RBD-1 demonstrated various abnormalities in the *C. elegans* development, such as defects in morphology (dumpy), incomplete molting, and defective gonadal and vulval development. Animals depleted for RBD-1 arrested mainly at the L1 larval stage.

In the course of studying the homeobox genes, we often used the dye-filling assay. It is the simplest method presently used to assay the structural integrity of sensory cilia. In paper II, we optimized conditions, in which reliable staining of the inner labial (IL2) neurons could be obtained, namely in low salt conditions, in the presence of detergent and when shaking worms at high speed during staining. The modified dye-filling procedure provides a reliable method to distinguish mutant alleles that stain amphids and phasmids, and IL2 neurons. Using this assay, we found that a mutation in the POU homeobox gene *unc-86* abolished dye-filling in IL2 neurons but not amphids and phasmids.

In Paper III, we showed that the LIM homeobox gene *ceh-14* was expressed in other sensory neurons and interneurons, including the phasmid neurons and the ALA interneuron, while previously it was shown that *ceh-14* is expressed in the AFD

neurons and required for thermotaxis behavior in *C. elegans*. *ceh-14* mutant animals displayed defects in dendrite outgrowth of the phasmid neurons, while the ALA interneuron and some tail neurons showed abnormal axonal outgrowth and pathfinding. *ceh-14* and the paired-like homeobox gene *ceh-17* act in the separate pathway to control normal axonal outgrowth of ALA neuron. Overexpression of CEH-14 in the nervous system may titrate out interacting factors, such as LDB-1, which caused developmental defects.

In paper IV, we investigated the function of four homeobox genes, *ceh-6, ceh-26, ttx-1* and *ceh-37*, in the excretory cell development. We showed that the POU-III class homeobox gene *ceh-6*, the Prospero class homeobox gene *ceh-26*, and two otd/Otx family homeobox genes, *ceh-37* and *ttx-1* formed a regulatory hierarchy required for the development and function of the excretory cell in *C. elegans*. The excretory cell is required for maintaining osmotic balance and excreting waste products. While *ceh-6* has previously been demonstrated to play a role in the excretory cell patterning, we showed here that *ceh-26* and *ceh-37* are expressed in the excretory cell. *ceh-26* mutants arrested in early larval development with defects characteristic for a lack of excretory cell function. Double mutant of the *otd/Otx* genes *ceh-37* and *ttx-1* was displaying larval arrest, consistent with the excretory cell dysfunction, which indicates that there is functional redundancy between these two genes. Using mutant alleles and RNAi, we showed that *ceh-26*::GFP and *ceh-37*::GFP was down-regulated in *ceh-6* mutants. Further, we found that *ceh-37*::GFP was down-regulated in the *ceh-26* mutants. We also examined the effect of the homeobox gene mutants on functional genes, such as channel proteins (the "target genes") that are expressed in the excretory cell and found that only a subset of the genes regulated by *ceh-6* was also regulated by *ceh-37/ttx-1*. We mapped the promoter regions of *ceh-26* and of the target gene *clh-4* to identify putative homeodomain proteins binding sites. Given that these homeobox genes are well conserved in evolution, we may expect that parts of this cascade are also conserved in other organisms.

LIST OF PUBLICATIONS

This thesis is based on the following papers and manuscripts, which are referred to in the text by their roman numerals:

I. Björk, P., Baurén, G., Jin, S., **Tong, Y. G.**, Bürglin, T. R., Hellman, U. and Wieslander, L. (2002). A novel conserved RNA-binding domain protein, RBD-1, is essential for ribosome biogenesis. *Mol Biol Cell* **13**,

II. **Tong, Y. G.** and Bürglin, T. R. (2010). Conditions for dye-filling of sensory neurons in Caenorhabditis elegans. *J Neurosci Methods* **188**, 58-61

III. Kagoshima H., Cassata G., **Tong Y. G.**, Pujol N., Niklaus G., and Bürglin, T. R. (2010). The LIM homeobox gene *ceh-14* regulates neurite outgrowth in select neurons. *Manuscript.*

IV. **Tong Y. G.**, Meemon K., Meemon K., Prétôt R., Viktorin-Aspöck G. and Bürglin, T. R. (2010). A regulatory network of homeobox genes is required for the function of the *Caenorhabditis elegans* excretory cell. *Manuscript.*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

4D	Three spatial dimensions over time
aat	Amino acid transporter
ABC	ATP-binding cassette
Antp	Antennapedia
Aqp	Aquaporin
Bcd	Bicoid
cdc	Cell division cycle
ceh	*C. elegans* homeobox gene
clh	CLC-type chloride channel
CNS	Central nervous system
cnx	Calnexin
cpb	Cytoplasmic polyadenylation element binding
CRX	Cone-rod homeobox gene
Ct	*Chironomus tentans*
daz	Deleted in azoospermia
dcp	Deacetylase complex protein
dfr	*drifter*
DIC	Differential interference contrast
DiD	1,1′-dioctadecyl-3,3,3′,3′-tetramethylindodicarbocyanine perchlorate
DiI	1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate
DiO	3,3′-dihexadecyloxacarbocyanine perchlorate
dis	Yeast disjunction abnormal
dyf	Abnormal dye filling
ELAV	Embryonic lethal, abnormal vision
EMSA	Electrophoretic mobility shift assay
epi	Abnormal epithelia
etr	ELAV-Type RNA binding-protein family
Ex	Excretory cell-specific *cis*-element
EXC	Excretory canal abnormal
FITC	Fluorescein isothiocyanate
fog	Feminization of germline
fox	Feminizing gene on X

HD	Homeodomain
HNF	Hepatocyte nuclear factor
IL	Inner labial
inx	Innexin
kin	Kinase
lag	*lin-12* and *glp-1* phenotype
lam	Laminin
ldb	LIM domain binding
let	Lethal
LHX	LIM homeobox gene
LIM	*Lin-11, isl-1, mec-3*
Lin	Abnormal cell lineage
LMO	LIM domain only
LMX	LIM homeobox gene
mec	Mechanosensory abnormality
mig	Abnormal cell migration
mls	Mesodermal Lineage Specification
msi	Musashi (fly neural)
nac	Na+-coupled dicarboxylate transporter
nas	Nematode astacin protease
nhr	Nuclear hormone receptor family
nhx	Na/H exchanger
NMJ	Neuromuscular junctions
NRE	Nuclear receptor binding element
NuRD	Nucleosome remodeling and deacetylating
Oct	Octamer-binding transcription factor
otd	*orthodenticle*
Otx	*orthodenticle* homeobox gene
pad	Patterning defective
pat	Paralysed arrest at two-fold
pgp	P-glycoprotein
POU	Pit-1, Oct-1, *unc-86*
PRD	Paired
Pros	Prospero

Prox	Prospero-related homeobox
QRT-PCR	Quantitative real time polymerase chain reaction
ral	Ras-related GTPase
RBD	RNA binding domain
Rh	Rhodopsin
RNAi	RNA interference
RNP	Ribonucleoprotein domain
RRM	RNA-recognition motif
sel	Suppressor/enhancer of *lin-12*
SIX	Sine oculis homeobox, Drosophila
sp	*Strongylocentrotus purpuratus*
spn	Spindle orientation defective
sulp	Sulfate permease
tag	Temporarily assigned gene
TALE	Three amino acid loop extension
ttx	Abnormal thermotaxis
unc	Uncoordinated
vha	Vacuolar H+-ATPase
VNC	Ventral nerve cord
vvl	Ventral veinless
ZF	Zinc finger

1 INTRODUCTION

The study of model organisms is based on the fact that all living species, which represent the diversity of life, descent from a common ancestor and therefore display conservation of metabolic and development pathways and genetic material over the course of evolution (Fox, 1985). Models are organisms that have plenty of biological data, and properties that make them attractive to study as examples and address questions, which are more difficult to study directly in higher organisms (including humans). These can be classified as genetic models, experimental models, and genomic models, with a pivotal position in the evolutionary tree. So far almost everything we know about the fundamental properties of living cells has come from the studies of model organisms (Fields and Johnston, 2005).

1.1 *Caenorhabditis elegans*

The fruit fly *Drosophila melanogaster* and the nematode worm *Caenorhabditis elegans* (**Figure 1**) are both important genetic models that have advantages such as short generation times, techniques for genetic manipulation, non-specialized living requirements, and rapid assays for phenotypic abnormalities. *D. melanogaster* was introduced and studied in the laboratory for the genetics work during the early 1900's (Kohler, 1994), and it has been used to address most of the questions and to obtain the present knowledge about the role of genes in the development of a multicellular animal. In the 60's, Sydney Brenner proposed the nematode worm *C. elegans* as another genetic model to investigate how various genes control embryonic development and the function of the whole nervous system (Brenner, 1974). The worm's simple body plan, a limited number of cells, a transparent body, and a relatively small genome (100 million base pairs, ~20 000 protein-coding genes) were considered its major advantages over the fruit fly (Epstein et al., 1995).

The genetic manipulation of *C. elegans* is easy. Its small size (~1mm), large number of progeny (300 to 1000 animals from a single hermaphrodite), fast generation time (~3.5 day), and ease of growth in either agar plates or in liquid cultures, allow for rapid growth and ease of cultivation in a laboratory. Therefore large mutagenesis screens can

be performed with *C. elegans*. Such screens are used to detect either specific phenotypes or deletions in known genes. Gene knockouts and genetic mutants can be

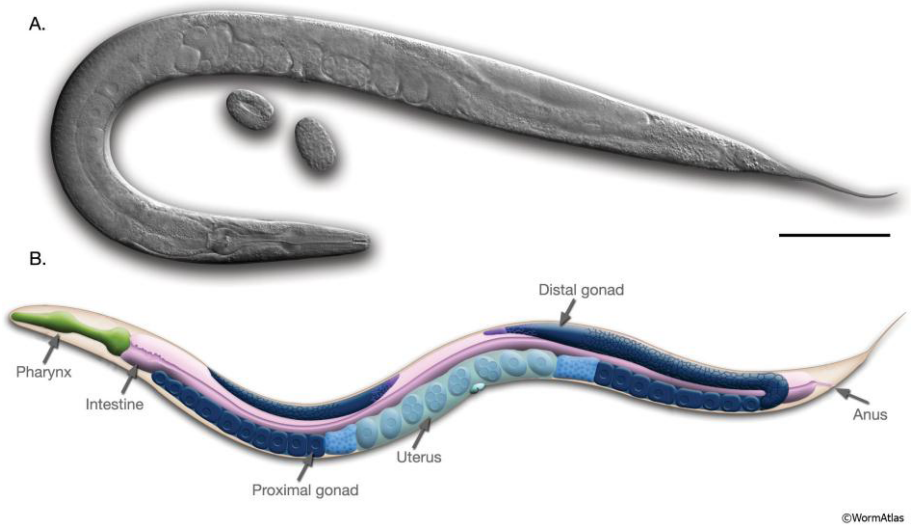

Figure 1 Anatomy of an adult hermaphrodite. A. DIC image of an adult hermaphrodite left lateral side. Scale bar 0.1 mm. B. Schematic drawing of anatomical structures, left lateral side. (Altun, Z.F. and Hall, D.H. 2009. Introduction. In www.wormatlas.org)

identified quickly, and gene knockdowns are readily generated via RNA-mediated interference (Fire et al., 1998). Males can be used for crossings to rapidly generate new genotype combinations. Transgenes can be introduced into the nematode either by microinjection or particle bombardment (Evans (ed.), 2006). However, insertion of DNA into defined regions of the genome by homologous recombination is not yet possible in *C. elegans*. All worm strains can endure freezing and thawing. Frozen stocks can be stored indefinitely in -80°C freezers or liquid nitrogen (−196°C) (Stiernagle, 2006).

As the worms are transparent, cells in the living animal can be observed under the differential interference contrast, DIC (Nomarski) microscopy, from the embryo to the adult worm. Reporter-gene fusions also allow direct visualization of cellular morphology and protein expression patterns. The lineage from the zygote to each of the 959 somatic nuclei of the adult hermaphrodite and the 1031 nuclei of the male has been identified. Such lineages are invariant between animals (Sulston, 1983; Sulston and

Horvitz, 1977; Sulston et al., 1983). Lineage analyses can be performed routinely with spatio-temporal (4D) microscopy in order to compare wild-type and mutant development (Hutter and Schnabel, 1994; Zipperlen et al., 2001).

Although the worm is tiny and simple, it is a multicellular organism with skin (cuticule), muscles, intestine, a nervous system, an excretory system and a reproductive system (de Bono, 2003; Rankin, 2002). There is already a wealth of anatomical data. For instance, there is an electron microscopic reconstruction of the whole *C. elegans* body (White, 1986). Also, the complete wiring diagram of the nervous system has been determined (White, 1986; Wood, 1988).

C. elegans is becoming widely used in the study of developmental control as a model animal for over 40 years. Its sequenced genome, and biosynthetic and metabolic pathways are highly conserved with vertebrates, including most of the known components involved in cellular development and the nervous system (Consortium, 1998). In 2002 the Nobel Prize in physiology or medicine was awarded to three key researchers in the worm field, Sydney Brenner, John Sulston, and H. Robert Horvitz. In 2006, Andrew Fire and Craig Mello were awarded the same prize for their discovery of RNA interference in *C. elegans*. Moreover in 2008, Martin Chalfie shared the Nobel Prize in Chemistry for his work with GFP in *C. elegans*. (http://nobelprize.org)

1.2 Homeobox genes

Both *Drosophila melanogaster*, as well as *C. elegans*, have been used to study genes involved in ontogenesis for many years. Important evidence that development was controlled by particular regulatory genes came from homeotic mutants in *Drosophila* (Gehring, 1994; Gehring, 1998; Lewis, 1978). The Nobel Prize in Physiology and Medicine was awarded jointly to Edward B. Lewis, Christiane Nüsslein-Volhard and Eric F. Wieschaus for their discoveries concerning the genetic control of early embryonic development in 1995 (http://nobelprize.org). These homeotic gene mutants result in transformations of body segments; parts of the body are transformed into other body parts. One of the extensively studied examples of these is the gene Antennapedia (Antp), where - as the name suggests - a dominant mutation that causes the

transformation of antennal structures to the corresponding (homologous) leg structures (**Figure 2**).

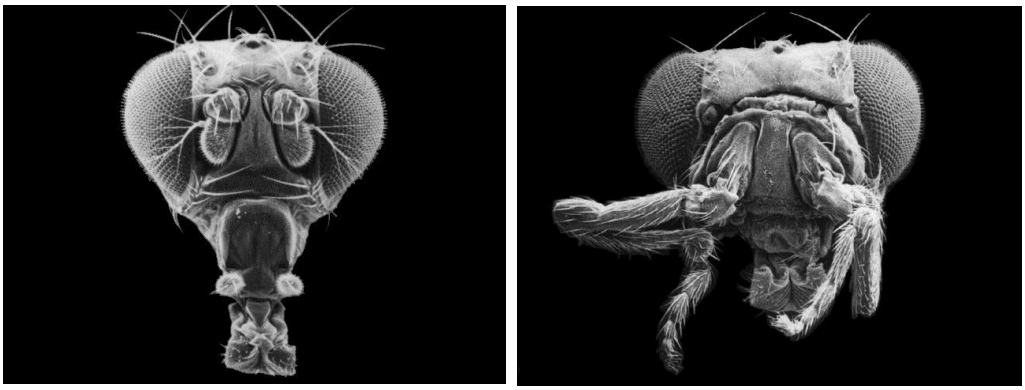

Figure 2 Frontal view of the fruit fly head. Wild type (left). Mutant Antennapedia (right). Antennae are transformed into metathoracic (second thoracic segment) legs (Turner and Mahowald, 1979).

When molecular techniques became available and were applied to cloning of these genes, it led to the discovery of the homeobox, a conserved stretch of DNA sequence of about 180 nucleotides that codes for the DNA-binding portion of a large class of transcription factors (McGinnis et al., 1984; Scott and Weiner, 1984). Subsequently, these genes were found to be involved in many aspects of development not only homeotic development. While "homeotic" was used to christen the "homeobox", it should be noted that not all homeotic genes are homeobox genes, and many homeobox genes do not have homeotic functions (Bürglin, 2005). Many more homeobox genes were subsequently identified from all eukaryotes. They have a wide phylogenetic distribution, having been found in baker's yeast, plants, and all animal phyla that have been examined so far (Derelle et al., 2007). Homeobox genes are essential for development as they control many different aspects of cellular growth and differentiation (Abate-Shen, 2002).

The homeobox encodes a DNA-binding domain of 60 amino acids that folds into three helices (**Figure 3**). The third helix is the key for providing specific DNA-binding (Qian et al., 1989). As hundreds of homeobox genes have been described, it became clear that the homeodomain of closely related genes have more sequence similarity. **Figure 4** shows a consensus sequence that is based on a compilation of 346 homeobox sequences. These 60 amino acid positions, especially the more conserved ones, define the typical homeodomain (Bürglin, 2005).

.

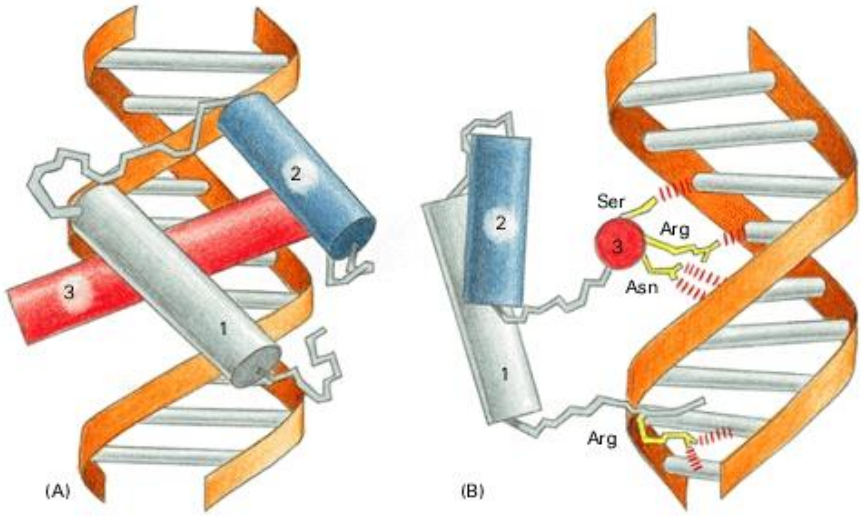

Figure 3 A homeodomain bound to its specific DNA sequence. The homeodomain is folded into three α helices, which are packed tightly together by hydrophobic interactions (A). The part containing helix 2 and 3 closely resembles the helix-turn-helix motif, with the recognition helix (red) making important contacts with the major groove (B). The Asn of helix 3, for example, contacts an adenine. Nucleotide pairs are also contacted in the minor groove by a flexible arm attached to helix 1. The homeodomain shown here is from a yeast gene regulatory protein (Alberts, 2002).

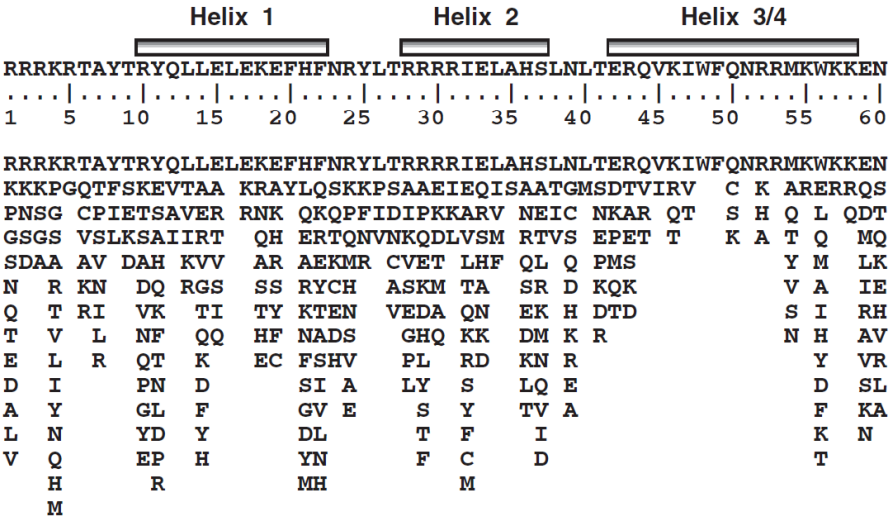

Figure 4 Homeodomain consensus sequence and amino acid encountered at a given position in the homeodomain (Bürglin, 2005).

Based on sequence similarities and fusion with characteristic co-domain sequences, homeodomain protein have been classified into several distinct classes, including TALE, Antp, PRD, SIX, LIM, POU, ZF, CUT, HNF and PRO (Bürglin, 2005; Booth and Holland, 2007; Takatori et al., 2008). The resulting classification allows to the examination of how structural conservation in sequence correlates with functional conservation. Many of the homeodomain classes have additional conserved protein domains or motifs outside of the homeodomain; some of the families are named after their extra domains (Bürglin, 2005).

The homeodomain can both bind DNA and mediate protein-protein interactions (Wolberger, 1996). Several sequences that are bound by homeodomain have been identified. For instance, the binding site of the homeodomain of the *Drosophila* morphogenic protein Bicoid (Bcd) is TAATCC. Bicoid is the only known protein that uses a homeodomain to regulate translation, as well as transcription, by binding to both RNA and DNA during early *Drosophila* development (Baird-Titus et al., 2006). POU homeodomain proteins usually bind to the octamer motif (ATGCAAAT) (Pierani et al., 1990; Staudt et al., 1986). Although the homeodomain in POU homeodomain proteins is essential for DNA binding, the POU-specific domain contributes specific contacts (Ingraham et al., 1990; Verrijzer et al., 1990). Therefore, the different types of homeodomain proteins target DNA in different ways to activate or repress transcription (or translation in some cases) of target genes.

Homeobox genes have been shown to play crucial roles from embryogenesis to cell differentiation in plants, fungi, and all animal phyla. It is important that one can compare the genomes of different organisms, since this will help us to understand the evolution of animal development. *C. elegans* genome encodes orthologs of most but not all developmental control genes known from *Drosophila* and vertebrates (Ruvkun and Hobert, 1998). Based on our data (Krishanu Mukherjee, unpublished), we identified 101 homeobox genes in *Caenorhabditis elegans*, 103 homeobox genes in *Drosophila melanogaster* and 339 homeobox genes in human.

1.2.1 OTD/OTX homeobox genes

Homeobox genes of the *orthodenticle (otd)/Otx* family belong to the class of paired-like homeodomain transcription factors that contain a lysine at position 50 of the homeodomain (Acampora et al., 2005; Acampora et al., 2001). In *Drosophila*, the only member of *otd/Otx* family is the *orthodenticle* (*otd*) transcription factor (synonym: *ocelliless*, *oc*). Its embryonic expression is essential for patterning of the anterior region of the body and the central nervous system (Finkelstein and Perrimon, 1990; Finkelstein et al., 1990; Hirth et al., 1995; Wieschaus et al., 1992; Younossi-Hartenstein et al., 1997). Severe or null mutant alleles are embryonic lethal, revealing defects in head structures, deletions in anterior parts of the brain, and defects in ventral nerve cords (Finkelstein et al., 1990; Klambt et al., 1991; Wieschaus et al., 1992). The viable mutant allele implicates *otd* in the terminal differentiation of the photoreceptor neurons of the fly eye (Ranade et al., 2008). In mice, three *otd*-related homeobox genes have been identified: *Otx1, Otx2* and *Otx5*/*CRX*. Both *Otx1* and *Otx2* are expressed anteriorly in the embryo in nested domains that include the embryonic forebrain and midbrain (Simeone et al., 1992). *Otx2* null mice die early in development due to severe defects in gastrulation and a lack of the anterior neuroectoderm that is fated to become forebrain, midbrain, and rostral hindbrain (Acampora et al., 1995; Ang et al., 1996; Matsuo et al., 1995). *Otx1* null mice are viable but have spontaneous epileptic seizures and abnormalities affecting the dorsal telencephalic cortex (Acampora et al., 1996).

All three transcription factors, *Otx1, Otx-2* and *CRX*, also contribute to the development of eyes (Ranade et al., 2008). In addition to the remarkable similarities in expression patterns and mutant phenotypes of the *otd/Otx* gene family, the *in vivo* cross-phylum rescue experiments provide further evidence for conservation of functional properties (Acampora et al., 1998; Nagao et al., 1998; Sharman and Brand, 1998). *otd/Otx* genes encode a homeodomain with a similar DNA-binding specificity as that of Bcd. This DNA-binding specificity is mediated by a lysine at position 50 of homeodomain (K50). In the known direct targets of Otd in the fly (Rh3, Rh5 and Rh6), gene transcription is regulated through TAATCC (GGATTA) sites located within the first few hundred base pairs upstream of the start of transcription (Tahayato et al., 2003). The sea urchin *otd*-related homeobox protein *SpOtx* is capable of binding the core consensus TAATCC to regulate the expression of the *Strongylocentrotus purpuratus Spec2a* gene (Mao et al., 1994).

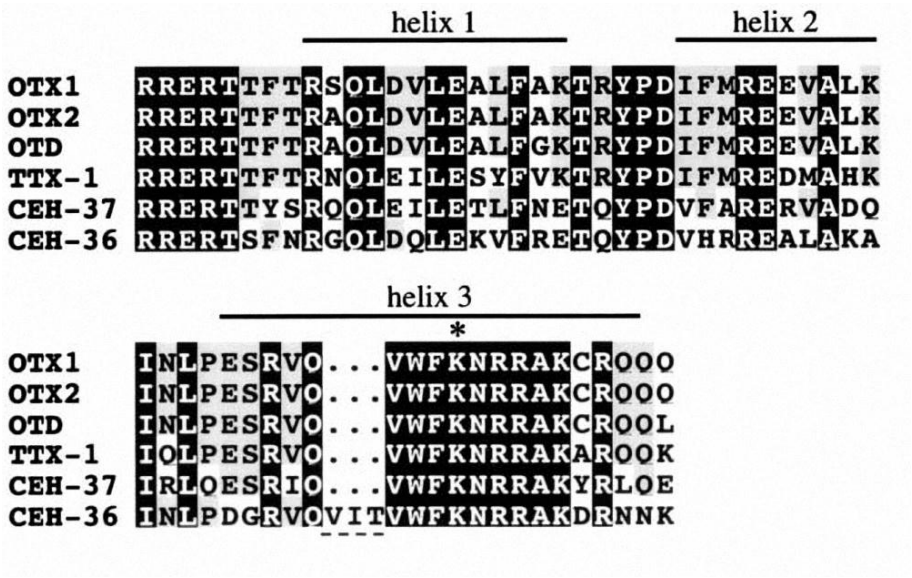

Figure 5 Alignment of the homeodomains of *C. elegans* CEH-36, CEH-37 and TTX-1, rat OTX1, rat OTX2, as well as *Drosophila* OTD. Identical residues are boxed in black; conserved residues are boxed in gray. The positions of the helices are over lined. The asterisk indicates the characteristic lysine at position 50. The three amino acid insertion in the third helix of CEH-36 is marked with a dashed underline (Lanjuin et al., 2003).

There are three *otd/Otx* genes in *C. elegans*: *ttx-1*, *ceh-37* and *ceh-36* (Galliot et al., 1999; Ruvkun and Hobert, 1998). The homeodomains of TTX-1, CEH-37 and CEH-36 show high conservation to the *otd/Otx* genes isolated from *Drosophila* and different vertebrate species, and the three genes contain the characteristic lysine at position 50 of homeodomain (Figure 5). However, the CEH-36 homeodomain shows a significantly lower homology with the other *otd/Otx* genes as its homeodomain contains an extra three-amino-acids insertion (VIT) in the predicted third helix (**Figure 5**). Phylogenetic analysis shows that CEH-36 and CEH-37 are more diverged from the vertebrate OTX proteins than TTX-1 (Lanjuin et al., 2003; Satterlee et al., 2001).

All members of OTD/OTX family of homeodomain proteins preferentially promote distinct neuronal identities in *C. elegans* (Koga and Ohshima, 2004; Lanjuin et al., 2003; Satterlee et al., 2001). *ttx-1* is expressed in the bilateral AFD thermosensory neurons and in the nine pharyngeal marginal cells that are structural cells of the pharynx with unknown functions. The AFD neurons have been assigned a

thermosensory role in *C. elegans* (Mori and Ohshima, 1995; Satterlee et al., 2001). The dendritic endings of the AFD neurons display a great increase in surface area and show finger-like shapes. *ttx-1* mutants show defects in the AFD neurons function. Electron microscopy of *ttx-1* mutants revealed that they lack dendritic endings (Perkins et al., 1986). Because *ttx*-1 regulates differentiated characteristics of the AFD neurons, *ttx-1* mutants are defective in a thermotaxis behavior and exhibit deregulated thermosensory inputs into a neuroendocrine signaling pathway. The AFD neurons also express partial AWC-like characteristics in *ttx-1* mutants. Ectopic expression of *ttx-1* converts other sensory neurons to an AFD-like fate. *ttx-1* expression in the AFD neurons was obviously reduced in *ttx-1* mutants , showing that TTX-1 autoregulates to maintain expression in the AFD neurons. In addition, TTX-1 regulates the expression of the LIM homeobox gene *ceh-14*, whose mutants exhibit weak thermostatic behavior (Cassata et al., 2000a; Satterlee et al., 2001). *ceh-37* is expressed in the AWB neurons and in the excretory cell. Mutations in *ceh-37* show defects in multiple aspects of AWB olfactory neuron specification. A GFP-tagged *ceh-36* transgene is expressed in the AWC and ASE neurons. *ceh-36* is required for the specification of AWC and ASEL (ASE in left side) neuronal identities. *ceh-36* is also required for establishing the left-right asymmetry in ASER chemosensory neurons (Chang et al., 2003). The HMX family homeodomain protein MLS-2 regulates *ceh-36* expression specifically in the AWC neurons. CEH-36 autoregulates itself by the K50-homeodomain binding site TAATCC to maintain its own expression (Kim et al., 2010). Moreover, CEH-36 may directly interact with the regulatory sequences of terminal differentiation genes in AWC and ASE neurons via the same binding motif (Chang et al., 2003; Etchberger et al., 2009; Lesch et al., 2009). All three *C. elegans otd/Otx* genes and rat *Otx1* can functonally substitute for *ceh-37* and *ceh-36*, but only *ceh-37* substitutes for *ttx-1*. Expression of both *ceh-37* and *ttx-1* in the AWB neurons restores AWB development, and TTX-1 compensates for CEH-37's function in the AWB neurons by activating the same downstream target genes (Lanjuin et al., 2003). Both *ttx-1* and *ceh-37* are also expressed in non-neuronal cell types in *C. elegans*, though the mutants do not show gross abnormalities in the development or function of these other types of cells.

1.2.2 POU homeobox genes

The POU class of homeobox genes contain the conserved POU-specific domain, which spans an approximately 80 amino acids, upstream of the homeodomain with a variable linker (Herr et al., 1988). The POU class of homeobox genes is further divided into five families (I, II, III, IV and V), which have diverse roles in cellular processes from stem cell functions to acting as transcription factors (Nichols et al., 1998; Ryan and Rosenfeld, 1997). *oct-4* (POU-V) is sufficient to reprogram mouse adult neural stem cells (Kim et al., 2009). In *C. elegans*, three POU homeobox genes can be identified: *unc-86* (POU-IV), *ceh-18* (POU-II) and *ceh-6* (POU-III). In *Drosophila* five genes in four families are found; no POU-1 family genes are identified in either flies or worms (Bürglin et al., 1989; Bürglin and Ruvkun, 2001).

unc-86 is required for the correct cell fate determination and differentiation in diverse neuronal lineages, including the egg-laying neurons, the mechanosensory neurons, and the chemosensory interneurons as well as all six IL2 neurons (Finney and Ruvkun, 1990; Finney et al., 1988). Although UNC-86 is not expressed in amphid sensory neurons, *unc-86* mutants are defective in their responses to chemosensory attractants mediated by the AWC and ASE sensory neurons (Bargmann et al., 1993). The AIZ interneurons, where *unc-86* is expressed, are probably important for chemotaxis (Mori and Ohshima, 1995). The POU member *ceh-18* plays a distinct role in development, specifying ectodermal and gonadal cell fates (Greenstein et al., 1994; Rose et al., 1997).

The third member of the POU homeodomain proteins in *C. elegans*, CEH-6 is expressed in particular head and ventral cord neurons, as well as in rectal epithelial cells, and in the excretory cell. *ceh-6* mutations result in embryonic lethality due to defects in the excretory cell function, that is required for osmoregulation. Expression of many genes in the excretory cell is abolished or downregulated in the *ceh-6* null mutant and in *ceh-6* RNAi knockdown animals (Armstrong and Chamberlin, 2010; Bürglin and Ruvkun, 2001; Mah et al., 2007; Mah et al., 2010). CEH-6 is also required for the function of several different types of neurons and epithelial cells (ref). The vertebrate POU-III genes (*Brn1*, *Brn2*, *SCIP/Tst1/Oct6/Pou3f1* and *Brn4*) are all expressed in the nervous system, particularly in the CNS (Alvarez-Bolado et al., 1995; He et al., 1989). Expression of POU-III genes has been detected in the excretory organs of mammals, quail, zebrafish, brine shrimp and *Xenopus* (Chavez et al., 1999; Lan et al., 2006;

Spaniol et al., 1996; Witta et al., 1995).

The optimal binding site of the POU- II (Oct-1 and Oct2) is the canonical octamer motif (ATGCAAAT) (Pierani et al., 1990; Staudt et al., 1986). Binding site of the *Brn* family (POU-III) is distinct from the *oct-1* motif (Wegner et al., 1993). *C. elegans* CEH-6 (POU-III) binds to the same *cis*-regulatory element (ATGCAAAT) to regulate the expression of *aqp-8* in the excretory cell (Mah et al., 2007). Mutagenesis assays indicate that the octamer *cis*-regulatory element could have a range of functional variants in *C. elegans* (Armstrong and Chamberlin, 2010).

1.2.3 LIM homeobox genes

LIM domain is about 60 amino acids long and made up of two contiguous zinc finger domains, separated by a two-amino acid residue hydrophobic linker (Kadrmas and Beckerle, 2004). They are named after their initial discovery in the proteins LIN-11, ISL-1 & MEC-3 (Bach, 2000). LIM-domains mediate protein interactions that are critical to cellular processes. LIM homeobox genes encode two LIM domains upstream of the homeodomain (Bürglin, 2005). This family of conserved genes plays a crucial role in neuronal survival, axon guidance and neuronal function. The diverse functions of LIM proteins include the development of kidney, pancreas, eyes, and limbs in vertebrates, the patterning of wings and imaginal disc precursor tissues in flies and the formation of the vulva in *C. elegans* (Hobert and Westphal, 2000).

The genome of the nematode *C. elegans* encodes seven LIM homeobox genes, which fall into the six subfamilies: apterous (*ttx-3*), LHX6/8 (*lim-4*), ISLET (*lim-7*), LMX (*lim-6*), LIM3 (*ceh-14*) and LIN1/5 (*mec-3* and *lin-11*). *ceh-14* is expressed in the thermosensory neurons AFDL/R, the interneurons BDUL/R and the asymmetric interneuron ALA. *ceh-14* mutants are athermotactic, but the AFD neurons are still present and differentiated (Cassata et al., 2000a).The LIM homeobox gene *ttx-3* is required for the functions of the interneuron AIY, including thermosensory behavior and olfactory learning (Altun-Gultekin et al., 2001; Remy and Hobert, 2005). *lin-11* plays a role in the differentiation of the AIZ interneurons (Hobert et al., 1998) which are a major component of the chemosensory and thermosensory circuits (White, 1986). These three LIM homeobox genes play a role in the *C. elegans*′ thermosensory

circuit, which contains AFD, AIY and AIZ neurons (Hedgecock and Russell, 1975).

The LIM domains are zinc-binding motifs interfacing with other proteins. These motifs do not only occur in homeodomain proteins, but also in other proteins that are involved in protein-protein interaction. In animals, there are LMO proteins (LIM domain only proteins), which have apparently secondarily lost the homeodomain. However, there are also more divergent LIM domains, which are found not only in eukaryotes, but also in bacterial proteins. They also function in protein-protein interaction (Bürglin, 2005; Duboule, 1994). The protein NLI/Ldb1/CLIM2 can bind to the LIM domains of the LIM homeodomain proteins but also to the nuclear LMO proteins to form functional complexes (Agulnick et al., 1996; Bach et al., 1997; Jurata et al., 1996; Milan and Cohen, 1999). Loss of LDB1 in mice causes severe pattering defects during gastrulation (Mukhopadhyay et al., 2003). *C. elegans* also contains an ortholog, *ldb-1*, which is expressed in all neurons and some other tissues of the larval and adult animal. The *ldb-1* RNAi worms exhibits the uncoordinated movement phenotype and defects in the function of mechanosensory and is thought to be an important co-factor for the LIM homeodomain proteins (Cassata et al., 2000b).

1.2.4 Prospero homeobox genes

Prospero (pros) class homeobox genes are atypical homeobox genes that contain a highly divergent 63-amino acid homeodomain. Three additional amino acids are located between helices 2 and 3. The conserved pros domain, which is about 100 amino acids long, lies downstream of the homeodomain. The pros domain forms a single structural unit with the homeodomain which provides sequence specific DNA binding (Holland et al., 2007). *Drosophila* pros is expressed in neuronal precursor cells and its asymmetric segregation determines the cell fates of sibling cells (Hirata et al., 1995; Spana and Doe, 1995). Mouse *Prox1* plays divergent roles in multiple internal organs (Burke and Oliver, 2002; Dudas et al., 2008; Sosa-Pineda et al., 2000). Its mutants are embryonic lethal, lacking the lymphatic vasculature around midgestation (Johnson et al., 2008; Wigle and Oliver, 1999). *Prox1* is also known to regulate the activity of a specific subset of nuclear receptors (Charest-Marcotte et al., 2010; Lee et al., 2009; Qin et al., 2004; Song et al., 2006; Yamazaki et al., 2009). Mouse *Prox2* is expressed in eyes, testes, as well as the developing nervous system. *Prox2* is not essential for

embryonic development and postnatal survival in mice (Nishijima and Ohtoshi, 2006). Human and chicken Prox1 mRNAs are detected in multiple organs including kidney in the embryonic and adult tissue (Rodriguez-Niedenfuhr et al., 2001; Zinovieva et al., 1996). Prox1 *in situ* hybridization of chicken embryos shows that a few tubuli of the kidneys are Prox1-positive (Rodriguez-Niedenfuhr et al., 2001). *Drosophila* Pros and human Prox1 have a unique sequence preference and recognize the following consensus sequence: AA/TC/TNNCT/C (Cui et al., 2004; Hassan et al., 1997). Only a single *prospero* class gene is present in *C. elegans*, *ceh-26*, which has been the subject of a few unpublished studies in the Bürglin laboratory.

1.3 RNA-binding proteins containing an RBD domain

The RBD (RNA binding domain), also known as RRM (RNA-recognition motif) or RNP (ribonucleoprotein domain), is one of the most common protein domains in eukaryotes and is extensively studied in terms of structure and biochemistry (Maris et al., 2005). A typical RBD is about 90 amino acids long with a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology that forms a four-stranded β-sheet packed against two α-helices (**Figure 6**) (Ding et al., 1999). RBD-containing proteins play a role in most post-transcriptional gene regulation processes (i.e. mRNA and rRNA processing, RNA export and stability) (Dreyfuss et al., 2002).

The RBD domain is not only involved in RNA/DNA recognition but also in protein-protein interaction. The structural versatility of the RBD interactions result in RBD-containing proteins having many diverse biological functions (Clery et al., 2008). Proteins have been found containing from one to several RBDs. The combination of two or more RBD domains allows the continuous recognition of a long nucleotide sequence and often increases the binding affinity (Maris et al., 2005). The individual RBD domains are separated by linker sequences of highly variable length which provide a critical determinant of binding affinity and may modulate *cis* versus *trans* binding (Shamoo et al., 1995). Several of proteins containing RBD domains have been studied in *C. elegans*: DAZ-1, SPN-4, FOG-1, CPB-1, RNP-4, FOX-1, UNC-75, EXC-7, ETR-1, MEC-8, MSI-1, EXC-7 and LIN-28 (Lee and Schedl, 2006). They are required not only in the germline and early development but also in the somatic development. The *C. elegans* neural RNA-binding protein ELAV homologue EXC-7,

which contains three RBD domains, is localized to nucleus in excretory cells and in the nervous system and regulates its mRNA targets (i.e. *sma-1* that encodes $ß_H$-spectrin) through post-transcriptional control, which further affects the proper development of the excretory canals (Fujita et al., 2003; Loria et al., 2003).

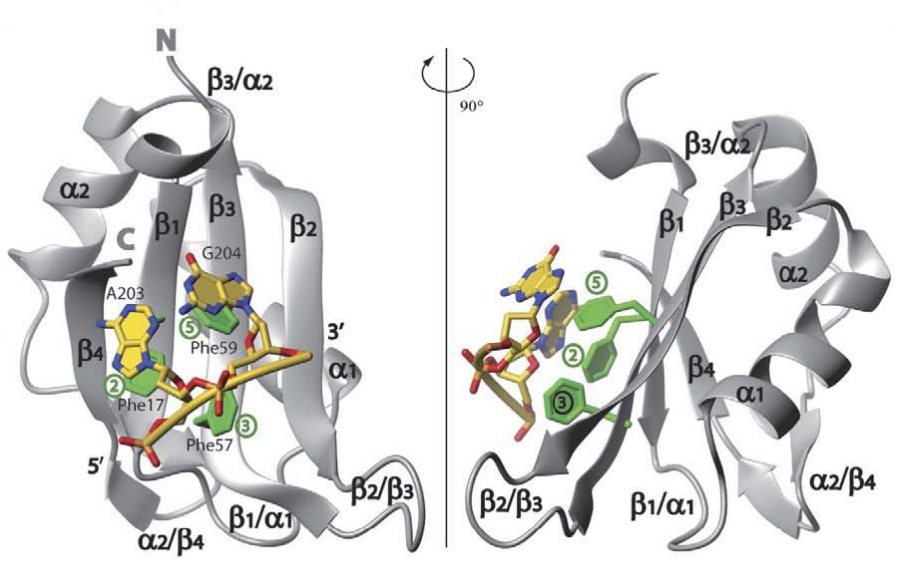

Figure 6 RRM (RBD) β-sheet interaction with RNA. Structure of hnRNP A1 RRM2 in complex with single stranded telomeric DNA as a model of single stranded nucleic acid binding. The place of main conserved RNP1 and RNP2 aromatic residues on the β-sheet indicated in green. Most commonly, three aromatic side-chains located in the conserved RNP1 (in β3-strand) and RNP2 (in β1-strand) sequence stretches, accommodate two nucleotides as follows: the bases of the 50 and of the 30 nucleotides stack on an aromatic ring located in β1 (position 2 of RNP2) and in β3 (position 5 of RNP1), respectively. The third aromatic ring that is usually located in β3 (position 3 of RNP1) is often inserted between the two sugar rings of the dinucleotide (Clery et al., 2008).

1.4 The excretory system of *C. elegans*

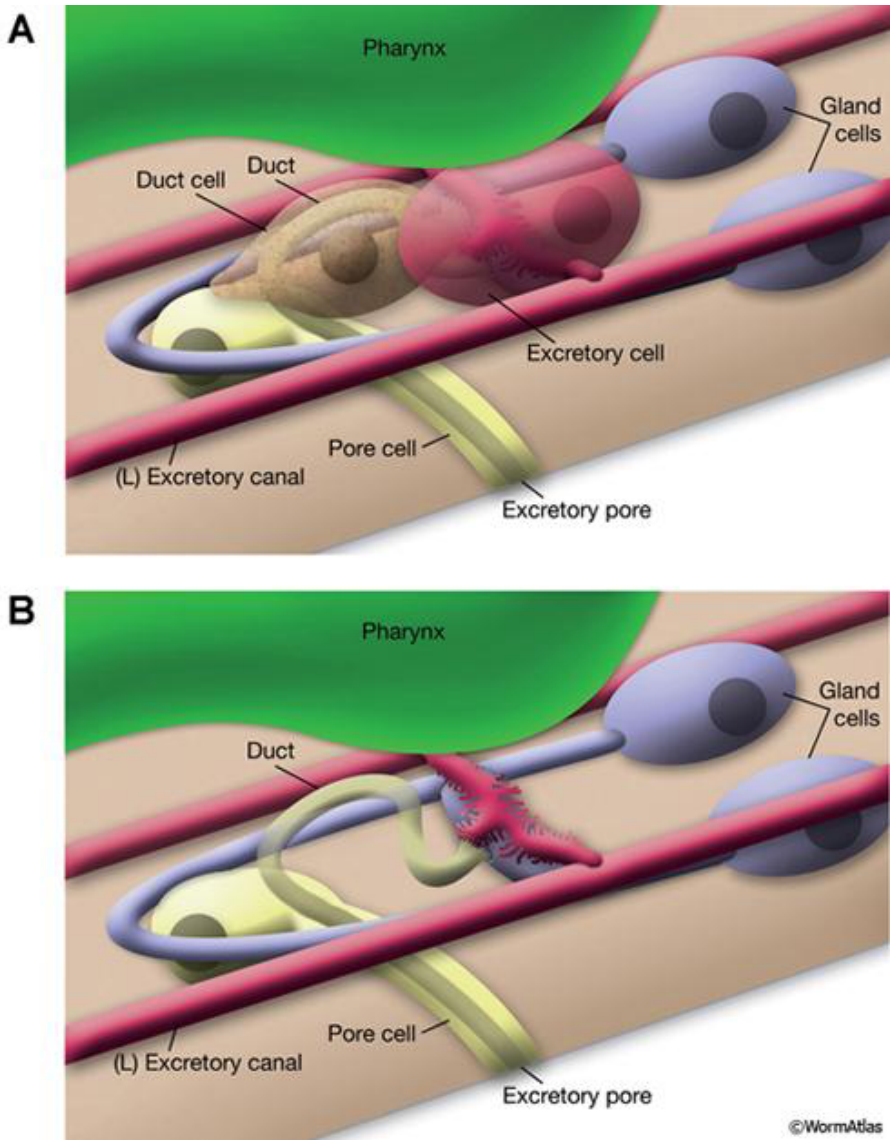

Figure 7 Schematic of the adult excretory system in *C. elegans.* Lateral oblique view. A) The fused pair of gland cells (*blue*), the excretory (canal) cell (*red*), the duct cell (*brown*) and the pore cell (*yellow*) are located in the head region. B) In this view, excretory and duct cell bodies are removed to reveal the junction. (Altun, Z.F. and Hall, D.H. 2009. Excretory system. In www.wormatlas.org)

The excretory system (**Figure 7**) in nematodes acts in osmotic and ionic regulation (Nelson et al., 1983; Nelson and Riddle, 1984) and is also responsible for secreting hormones and fluids required for molting (Nelson et al., 1983). This system consists of

four cell types: one excretory duct cell, two excretory gland cells, one excretory pore cell, and one excretory (canal) cell in *C. elegans*. All cells are descendants of the AB cell lineage, but no direct precursor cell is shared amongst these cells (Sulston and Horvitz, 1977). Laser ablation of any of the three central cells, i.e. the excretory cell, the duct cell, and the pore cell, leads to fluid retention in the worm (Nelson and Riddle, 1984). These cells and their morphology are highly variable in nematodes (Chitwood and Chitwood, 1977; Wang and Chamberlin, 2002). *C. elegans* exhibits the greatest ability to survive in a high salt test among five different Caenorhabditis species (*C. elegans*, *C. remanei*, *C. brenneri*, *C. briggsae* and *C. japonica*) (Wang and Chamberlin, 2004).

1.4.1 Excretory cell

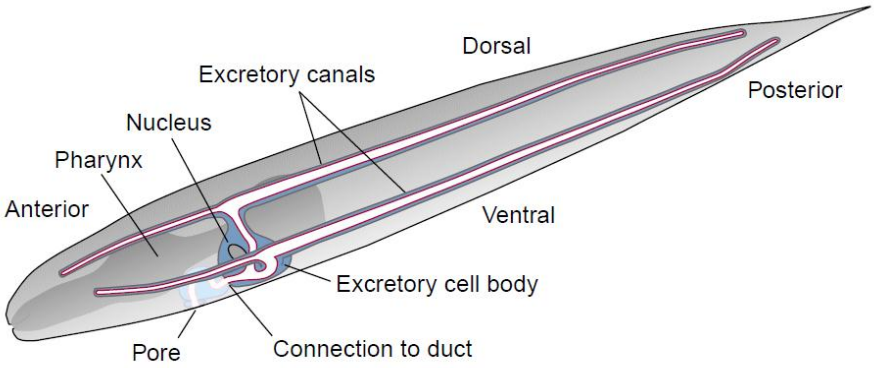

Figure 8 Structure of the excretory canals in a late larval stage worm. The apical/luminal (red) and basal (blue) surfaces of the excretory cell meet only at the connection to the duct cell (Buechner, 2002).

The H-shaped excretory cell (**Figure 8**) is the largest cell in *C. elegans* and its birth at the end of gastrulation is a useful identifiable landmark in embryogenesis (Fujita et al., 2003). The excretory canals are fluid-filled channels surrounded by the cytoplasm of the excretory cell processes, which extend both anteriorly and posteriorly from the excretory cell body beneath the hypodermis on the worm's lateral side. The surface areas of canals are increased by the structure of canaliculi (**Figure 9**), which also

provide exposure to the extracellular fluid-filled pseudocoelomic cavity (Buechner et al., 1999).

Most tubules are made of a series of linked epithelial cells. As single-celled tubules, the excretory cell of the nematode *C. elegans* is a good example of cell morphogenesis but also a simply genetically traceable model for studying tubule formation and regulation, just like the other model, the single-cell tips of the *Drosophila* trachea (Buechner, 2002; Casanova, 2007).

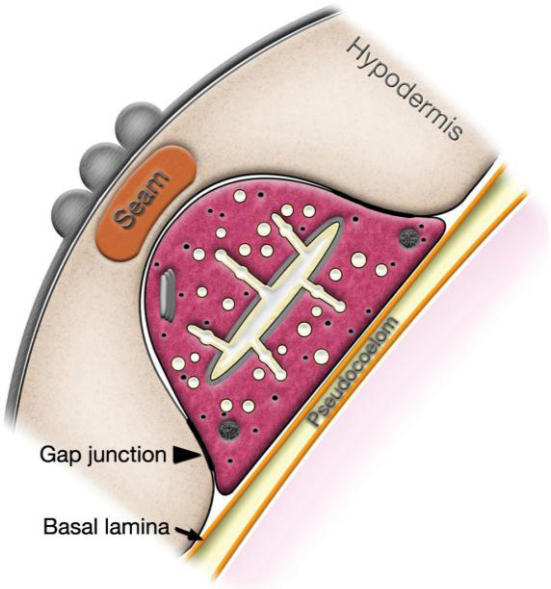

Figure 9 Illustration of an excretory cell canal cross section. Excretory cell tubes locate an apical membrane that forms the lumen. Numerous canaliculi branch off of the center lumen (Altun, Z.F. and Hall, D.H. 2009. Excretory system. In www.wormatlas.org)

The trachea of *Drosophila* is a network of tubes that fulfills the respiratory needs of the fly. Several homeobox genes including *Antennapedia*, *abdominal*-A, *abdominal*-B, *cut, engrailed, empty spirades, nubbin* and *ventral veins lacking* are found in tracheae and spiracles (http://www.sdbonline.org/fly/aimorph/trachia.htm). The fly gene *ventral veinless* (*vvl*) which is ortholog of the *C. elegans* gene *ceh-6* (POU-III), corresponds to the genetic locus *drifter* (*dfr*), also known as *Cf1a,* which regulates the developmental processes of tracheal formation (Anderson et al., 1995; de Celis et al., 1995; Llimargas and Casanova, 1997). However, *dfr* is not expressed in the *Drosophila* malpighian tubules, which carry out the function of waste excretion and maintenance of osmotic homeostasis, like the *C. elegans* excretory cell (Jung et al., 2005). So far only three homeobox genes (*caudal, cut,* and *homothorax*) are detected in the fruit fly *Drosophila melanogaster* malpighian tubules (http://www.sdbonline.org/fly/aimorph/maltubls.htm).

This suggests that a different homeobox genes transcription regulation network may be active in excretory organs of *C. elegans* and the *Drosophila*.

Due to some similarities in their structures (e.g., outgrowth of processes), the excretory cell and neurons share many developmental characteristics that direct elongation, guidance and outgrowth. Indeed, mutations in many genes, including *pat-3* (β_1-integrin), *epi-1* (the laminin α chain), *lam-1* (laminin β chain), *unc-52* (perlecan), *mig-15* (Nck-interacting kinase), *unc-53* (a novel actin-binding protein), *unc-73* (a rho-type GDP–GTP exchange factor), *mig-10* (Grb7), *unc-34* (a EVH1 domain-containing protein), *unc-104* (kinesin), *unc-116* (kinesin), *un*c-6 (netrin), *unc-5* (a netrin receptor) and *lin-17* (a Frizzled receptor), affect the development of both excretory cell and neurons in elongation, guidance, outgrowth, and cell fate decision. However, none of the above mentioned mutations causes defects in the ability of the canals to form tubules or to regulate the diameter of the lumen (Buechner, 2002). In addition, excretory cell tubulogenesis occurs *in vitro* using conditions intended for neuronal cell culture (Buechner et al., 1999).

1.4.2 EXC genes regulation pathway

Buechner *et al* in 1999 first reported a number of the lumenal mutants in the excretory cell of *C. elegans*. These mutants are unable to maintain a narrow channel structure and would swell into a series of fluid-filled cysts (EXC), in the excretory cell of *C. elegans*. So far, mutations causing a visible EXC phenotype have been found in twelve genes including *exc-1, exc-2, exc-3, exc-4, exc-5, exc-6, exc-7, exc-8, exc-9, let-4, let-653* and *sma-1* (Buechner et al., 1999). EXC proteins act together to regulate the excretory channel shapes (**Figure 10**). The model of EXC protein function suggests that *sma-1* mRNA, which encodes a β_H –spectrin is transported via EXC-7 ELAV, which encodes a RNA binding domain-containing protein (see above), and is translated along the length of the canals (Tong and Buechner, 2008). EXC-9, which encodes a LIM domain, activates EXC-1 and EXC-5 (encodes a guanine nucleotide exchange factor) to cause CDC-42 (encodes a RHO GTPase) to be phosphorylated, and direct the growth and stability of the apical cytoskeleton, while inhibiting the basolateral cytoskeleton. In this model, a gap at the apical cytoskeleton probably triggers activation of EXC-9, while the spectrin at the apical surface would repress activation of EXC-9. *exc-4* encodes a CLC

chloride channel (Berry et al., 2003) and its expression is necessary for EXC-9 to function normally. *exc-2* and *exc-1* genes are not yet cloned, but mutants analysis indicates that both the presence of the EXC-4 CLC channel and EXC-2 allow EXC-9 to function and EXC-1 may function together with EXC-5 downstream of EXC-9 in a pathway that activates CDC-42 at the canal apical surface (Tong and Buechner, 2008). Mutations in *let-4* and *let-653* result in an extremely large canal lumen which is visible under the dissecting microscope. *let-4* has not yet been cloned, and *let-653* encodes a secreted mucin (Jones and Baillie, 1995). Its functional pathways are not yet known.

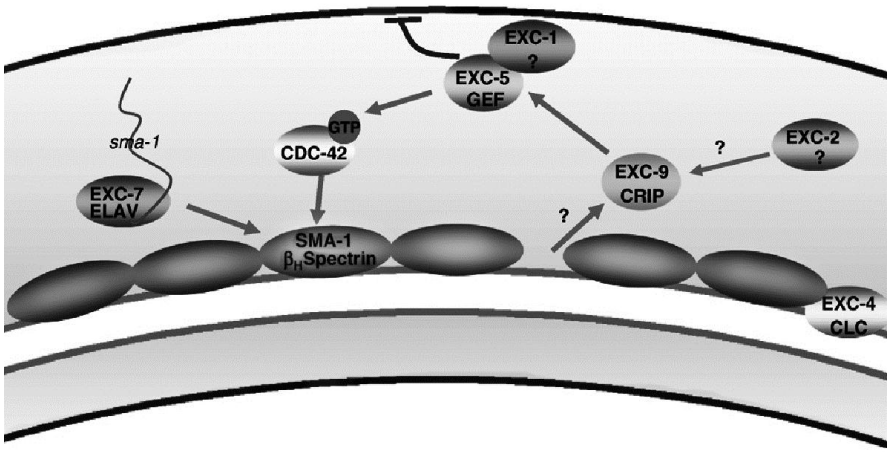

Figure 10 Model of EXC protein interactions. Diagram depicts a section of the excretory canals (Tong and Buechner, 2008).

1.4.3 CEH-6 dependent genes

The EXC phenotypes or cysts in the excretory cell are not only seen in the mutants discussed above. The homeobox gene *ceh-6* (POU-III family) mutants display similar cysts in the excretory channel and/or body of *C. elegans*, and about 80% of the animal die in larval stages (Bürglin and Ruvkun, 2001). The transcription factor *ceh-6* is the only POU homeobox gene that is expressed in the excretory cell. It directly regulates the expression of *aqp-8*, which encodes an aquaglyceroporin and *nac-2*, which encodes a sodium-coupled dicarboxylate transporter, in the excretory cell via the octamer elements (ATTTGCAT) of their promoters (Armstrong and Chamberlin, 2010; Mah et al., 2007). *ceh-6* RNAi and mutants screens uncovered several of additional downstream genes, which are regulated in the excretory cell by CEH-6, but also

Table 1 Excretory cell gene: gene tested for CEH-6 dependence. Adapted from (Armstrong and Chamberlin, 2010; Bürglin and Ruvkun, 2001; Mah et al., 2007)

Gene	Description	CEH-6 dependence	Octamer element	Test
nac-2	Na^+ dicarboxylate transporter	Yes	Yes	Mutant and RNAi
clh-4	ClC-type chloride channel	Yes	Yes	RNAi
twk-36	K^+ channel transporter	Yes	Yes	RNAi
sulp-4	Sulfate permease anion transporter	Yes	Yes	RNAi
sulp-5	Sulfate permease anion transporter	Yes	Yes	RNAi
pgp-3	ATP-dependent transporter	Yes	Yes	RNAi
aqp-8	Water channel	Yes	Yes	RNAi
Y70G10A.3	Na^+ dicarboxylate transporter	Yes	Yes	RNAi
nhx-9	Na^+/proton exchanger	Yes	Yes	RNAi
sulp-8	Sulfate permease anion transporter	Yes	Yes	RNAi
vha-12	Vacuolar-type ATPase	Yes	Yes	RNAi
mrp-2	ATP-binding cassette (ABC) transport	Yes	Yes	RNAi
sel-9	p24 family of proteins	Yes	Yes	RNAi
ral-1	Ras-related GTPase	Yes	Yes	RNAi
aat-1	amino acid transporter	Yes	Yes	RNAi
pgp-4	ATP-binding cassette (ABC) transporter	Yes	No	RNAi
vha-13	Vacuolar-type ATPase	Yes	Yes	RNAi
kin-18	Serine/threonine protein kinase	No	Yes	RNAi
unc-115	Actin-binding LIM Zn-finger protein	No	Yes	RNAi
cnx-1	calnexin	No	Yes	RNAi
inx-13	innexin	No	Yes	RNAi
vha-1	Vacuolar-type ATPase	No	No	Mutant and RNAi
vha-5	Vacuolar-type ATPase	No	Yes	RNAi
vha-8	Vacuolar-type ATPase	No	Yes	RNAi
exc-9	LIM domain-containing proteins	No	Yes	RNAi

showed that some genes are CEH-6 independent (**Table 1**). CEH-6 dependent genes that are found in **Table 1** are all terminal differentiation genes such as structural proteins, channels, and transporters, i.e. they are the targets of the regulatory transcription factors. The data summarized in Table 1 suggest that *ceh-6* play a role in the function of the excretory cell.

ceh-6(RNAi) does not suppress the expression of EXC-9 (Armstrong and Chamberlin, 2010) and therefore *ceh-6* may not be involved in the EXC protein interaction pathway. To date, we can find about 300 genes expressed in the excretory cell including a couple of transcription factors; for example, *ceh-37, ceh-26, tag-97, pad-1, nhr-31, nhr-17, nhr-41, nhr-43, nhr-91, dis-3* and *set-17* (http://www.wormbase.org). For comparison, the POU homeobox gene *unc-86* by itself allows activation of the expressions of the *mec-3*, a LIM homeobox gene, in the mechanosensory neurons (Lichtsteiner and Tjian, 1995). Does *ceh-6* have such a key regulatory role in the excretory cell? Future experiments will be necessary to determine, whether expression of the above mentioned transcription factors are dependent on CEH-6 regulation. Some interactions have been studied in this thesis.

1.4.4 Nuclear receptor and its binding elements (NREs)

The vacuolar ATPase (vATPase) is an ATP-dependent proton pump that transports protons across cellular membranes (**Figure 11**). Each *C. elegans vha* gene encodes one subunit of the vATPase holoenzyme, and there are 15 separate subunits that make up the holoenzyme (Hahn-Windgassen and Van Gilst, 2009). *ceh-6* RNAi knock-down down-regulates the expression of *vah-12* and *vah-13, but* has no effect on the expression of *vah-1*, *vah-5* and *vah-8* (**Table 1**), while RNAi of *nhr-31* dramatically affects the expression of the 15 genes that encode subunits of the vacuolar ATPase (**Figure 11**).

NHR-31 encodes one of 268 *C. elegans* nuclear receptors and is highly conserved to the mammalian and *Drosophila* HNF4 receptors (Robinson-Rechavi et al., 2005). Its expression is required for controlling the growth and function of the nematode excretory cell (Hahn-Windgassen and Van Gilst, 2009). RNAi of the vacuolar ATPase

subunits such as *vha-5* (small a subunit), *vha-8* (catalytic E subunit) or *vha-12* (B subunit) leads to tube formation phenotypes similar to those of *nhr-31* RNAi animals.

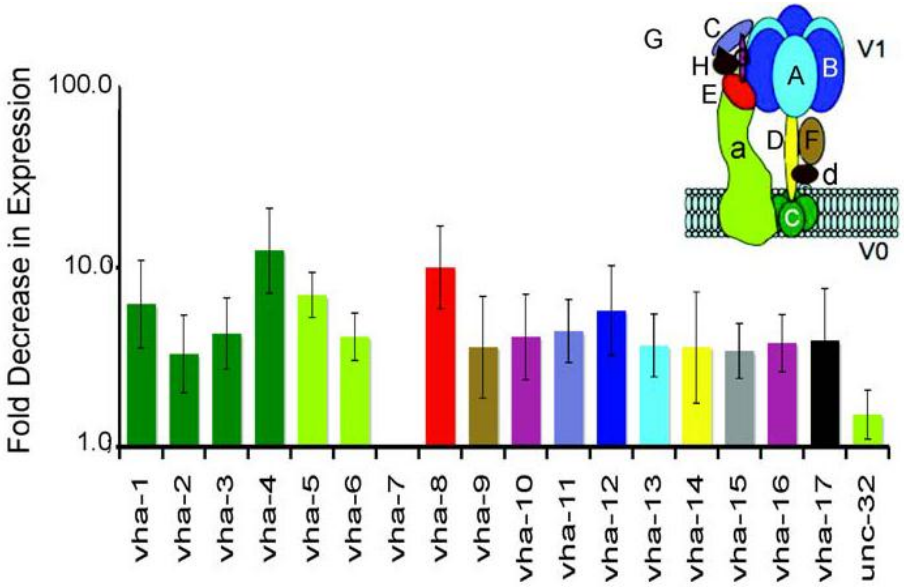

Figure 11 NHR-31 promotes expression of numerous genes encoding subunits of the vacuolar ATPase. The change in the expression levels of vacuolar ATPase genes (*vha* genes) in *nhr-31* RNAi animals is shown. Data were obtained by QRT-PCR and are plotted as the fold-decrease in gene expression observed in *nhr-31* RNAi animals as compared to animals fed control RNAi. The data are presented on a log scale and are color-coded according to the specific subunit encoded by that gene (the vATPase holoenzyme is shown in the inset, the peripheral domain is named as the V1 domain and the integral membrane domain is termed V0). For some subunits, several different gene isoforms exist. Error bars represent standard error (n=7). (Hahn-Windgassen and Van Gilst, 2009)

Nuclear receptors typically associate with complex binding motifs that comprised of two hexameric half-sites (Mangelsdorf et al., 1995). Nuclear receptor binding elements (NREs) are predicted in the promoters of 15 out of 18 *vha* genes (Hahn-Windgassen and Van Gilst, 2009), including *vah-12* and *vah-13* which are also regulated by *ceh-6* via the octamer element (Armstrong and Chamberlin, 2010).

1.4.5 Excretory cell-specific *cis*-element: Ex-1, Ex-2 and Ex-3

Figure 12 Alignment of the 500-bp *C. elegans* (CE) *pgp-12* and its 500-bp *C. briggsae* (CB) orthologous region using NCBI BLAST 2. The 10-bp element, Ex-1 (box), Ex-L (line), Ex-R (dashed line), and Ex-R1 (dots) are highlighted (Zhao et al., 2005).

Except for the octamer *cis*-regulatory elements discussed above, three other motifs, Ex-1, Ex-2 and Ex-3, have been identified in genes expressed in the excretory cell in *C. elegans* (Zhao et al., 2005). The 10 bp EX-1 (CCATACATTA) motif at -249 of the promoter of *pgp-12* was identified using deletional analyses of *pgp-12*-promoter-GFP fusion constructs (**Figure 12**). *pgp-12* encodes a P-glycoprotein subclass of ATP-binding cassette (ABC) transporters and it is expressed only in the excretory cell. *pgp-12* expression is regulated by DCP-*66*, which binds the *pgp-12* Ex-1 promoter element in yeast one-hybrid and EMSA experiments (Zhao et al., 2005). *dcp-66* encodes an ortholog of the NuRD component p66 and is a member of the transcriptional inhibitory nucleosomal remodeling and deacetylase (NuRD) complex. Human and *Xenopus* homologs of DCP-66 were identified as a component of a gene silencing complex (Feng et al., 2002; Wade et al., 1999). Similarly, it negatively regulates *lag-2*, which encodes a transmembrane protein of the Delta/Serrate/Lag-2 (DSL) family, and is expressed in the gut of *C. elegans* (Poulin et al., 2005). However, *dcp*-66 specifically up-regulates *pgp-12* expression in the excretory cell (Zhao et al., 2005). Deletion of the Ex-1 motif eliminates *pgp-12*::GFP expression at all developmental stages, while deletion of Ex-L, Ex-R or Ex-R1 (**Figure 12**) only partially lowers *pgp-12* expression in the excretory cell (Zhao et al., 2005). Since *dcp-66* does not seem to be a specific transcription factor, but rather a part of the NuRD complex, this apparently specific function seems surprising. In fact, there is an octamer-like motif ATTTCCAT within the Ex-1 and Ex-L elements and this suggests that CEH-6 can possibly regulate *pgp-12* expression in some unknown DCP-66 dependent way. A 26 bp (Ex-2) and a 25 bp element (Ex-3) have been identified in the promoters of *pgp-3* and *nas-31*, respectively,

with the following sequences, CTCACAAAATATAAATATGGTAATTC and GTTTCGAAAGTTCATCACCCCCAAC. Similar strategies as those used for *pgp-12* were used. *dcp-66* weakly reduces the expression of *pgp-3* and *nas-31* in the excretory cell and it suggests potential co-factors and cross-talk between transcription factors (Zhao et al., 2005).

1.5 The nervous system of *C. elegans* and Dye filling staining

The adult *C. elegans* hermaphrodite has 302 neurons (Sulston and Horvitz, 1977). They communicate via approximately 6400 chemical synapses, 900 gap junctions, and 1500 neuromuscular junctions (NMJs). Among individual animals, the location of the chemical synapses is about 75% reproducible (Durbin, 1987). Almost all *C. elegans* neurons have simple monopolar or bipolar structures. They fall into four functional categories defined by their circuitry: motor neurons, sensory neurons, interneurons and polymodal neurons which perform more than one of these functional characteristics. Except for these categories, there is a small subset of neurons whose functions are yet unknown (Hall et al., 2006). The invariant anatomy and the simple neuronal morphology make it very easy to detect even minor developmental defects and this makes *C. elegans* one of the favorite model organisms to study neuronal development.

C. elegans can detect and differentiate a wide range of chemical compounds, including water-soluble chemicals such as anions, cyclic nucleotides, cations, biotin, and amino acids and volatile chemicals such as alcohols, aldehydes, esters, ketones, pyrazines, thiazoles, and aromatic compounds (Bargmann, 2006). It senses chemicals with chemosensory neurons that penetrate the cuticle to expose their sensory cilia to the environment: the amphid, phasmid, and inner labial neurons (Ward et al., 1975; Ware, 1975). Of these neurons, 22 are paired amphid neurons (**Figure 13**) including ASE, ASG, ASH, ASI, ASJ, ASK, ADF, ADL, AWA, AWB, and AWC, four are paired phasmid neurons (PHA and PHB), and six are IL2 neurons. To date, the functions of these sensory neurons have been studied and identified, except for the inner labial neurons (IL2) (Bae and Barr, 2008; Bargmann, 2006). Fluorescent dyes such as FITC, DiI, DiO and DiD can fill the amphid chemosensory neurons in the head and the phasmid chemosensory neurons in the tail through their exposed ciliated endings, when living worms are placed in a solution containing such dyes (Hedgecock et al., 1985;

Shaham (ed.), 2006; Starich et al., 1995). FITC is able to label amphid neurons (ADF, ASH, ASI, ASK and ADL) and phasmid neurons (PHA and PHB) (Hedgecock et al., 1985). The lipophilic tracer dyes DiI, DiO and DiD can label the amphid neurons ASI, ADL, ASK, AWB, ASH and ASJ, as well as the phasmid neurons PHA and PHB (Coburn et al., 1998; Collet et al., 1998).

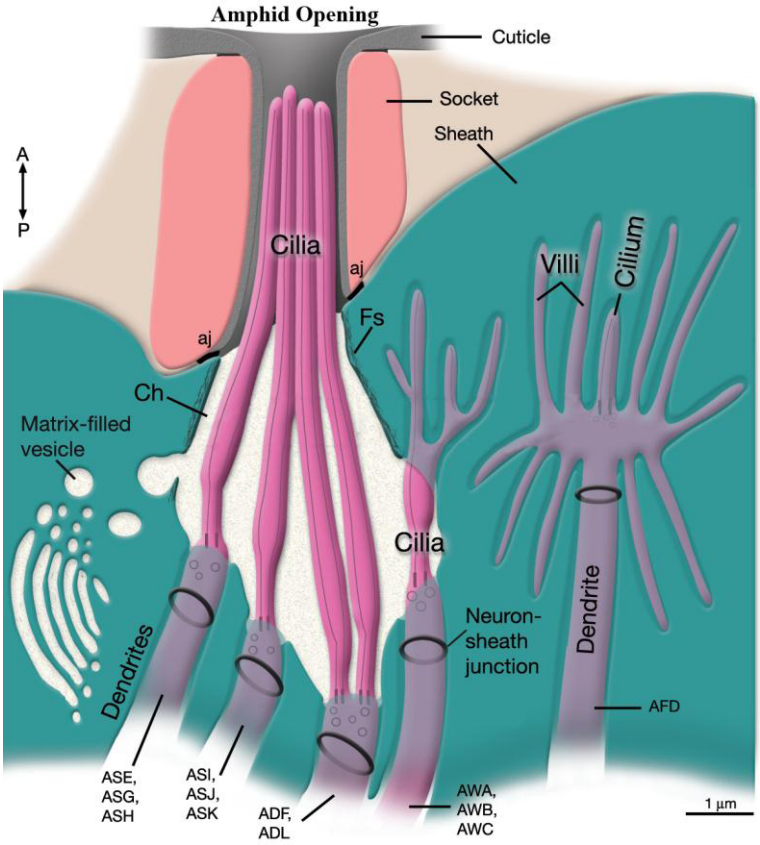

Figure 13 Structure of the amphid opening, longitudinal section. Amphid channel (Ch) is lined by the cuticle in the distal (socket) part and an electron-dense lining supported by a scaffold of cytoskeletal filaments (Fs) in the anterior sheath part. The socket cell is connected to the hypodermis and the sheath cell by adherens junctions (aj). Adherens junctions are also seen between the dendrites and sheath cell (neuron-sheath function) proximally to the level where the dendrites enter the channel. A large Golgi apparatus located at the base of the sheath-cell process (*left*) gives rise to matrix-filled vesicles bound toward the channel. Mitochondria (not shown) are also present in this region. (Altun Z.F. and Hall, D.H. 2009. Nervous system, neuronal support cells. In www.wormatlas.org)

Burket *et al*. developed a new dye-filling protocol for labeling the IL2 neurons and showed that it is possible to label the IL2 neurons with dye (Burket et al., 2006). So far,

the mechanism how dye is taken up is not fully understood. However, it is necessary that the ciliated neuronal processes are exposed to the environment so that the dyes can enter the neurons (Bae and Barr, 2008). Any defects in the ultrastructure of the chemosensory cilia or their accessory cells will prevent dye uptake and disrupt chemosensory behaviors (Perkins et al., 1986). Such dye-filling experiments are an efficient and simple method used to assay the morphological intactness of the cilia and sensory neurons (Inglis et al., 2007)

2 AIMS

The general aim of the study was to get a better understanding of the role of developmental regulators using the model organism *C. elegans.*

Specifically, the aims were to:

1. Investigate the role of a novel conserved RNA-binding Domain Protein.

2. Develop new dye-filling protocol in order to improve IL2 neuron identification and improve the detection of DYF phenotypes.

3. Study the function of different homeobox genes in development control and cell differentiation.

4. Define and characterize a regulatory network among several homeobox genes in the excretory cell

5. Investigate the evolutionary conversation of *cis*-regulatory elements of homeobox genes in the excretory cell

3 RESULTS AND DISCUSSION

3.1 PAPER I: A NOVEL CONSERVED RNA-BINDING DOMAIN PROTEIN, RBD-1, IS ESSENTIAL FOR RIBOSOME BIOGENESIS

The RBD is one of the most common protein domains in eukaryotes and genes encoding this domain play a role in wide variety of post-transcriptional gene regulation processes. In this paper we have characterized a novel protein, RBD-1 (RNA Binding-Domain-1), which is involved in ribosome biogenesis. The protein contains six conserved RNA-binding domains and is highly conserved from yeast to human. Biochemical studies in *Chironomus tentans* show that RBD-1, binds pre-rRNA in vitro and anti-Ct- RBD-1 antibodies repress pre-rRNA processing in vivo. Ct-RBD-1 is mainly located in the nucleolus in an RNA polymerase I transcription dependent manner, but is also present in discrete foci in the interchromatin and the cytoplasm. In cytoplasmic extracts, 20-30% of Ct-RBD-1 is associated with ribosomes and, preferentially with the 40S ribosomal subunit. cDNA clone yk417f6 (kind gift of Yuji Kohara) spans the whole open reading frame of *C. elegans rbd-1*. We have performed RNAi experiments using this cDNA and shown various abnormalities such as dumpy, incomplete molting, and incomplete and defective gonadal and vulval development. Animals arrest mainly at the L1 larval stage. Moreover, this gene is essential in *Caenorhabditis elegans*. Our data suggest that RBD-1 plays a role in structurally coordinating pre-rRNA during ribosome biogenesis and that this function is conserved in all eukaryotes.

3.2 PAPER II: CONDITIONS FOR DYE-FILLING OF SENSORY NEURONS IN *CAENORHABDITIS ELEGANS*

Dye-filling experiments are an efficient and simple method used to fast assay the morphological intactness of the cilia of a subset chemosensory neurons and defects of neuron support cells in *vivo* (Inglis et al., 2007). While the amphids and phasmids are easy to be stained, the six IL2 neurons are difficult to satin reproducibly. We optimized the conditions under which the IL2 neurons take up the lipophilic fluorescent dye DiI. In this paper, we describe that IL2 dye-filling depends on salt concentration, but not

osmolarity. The modified dye-filling procedure provides a reliable method to distinguish mutant alleles that stain amphids and phasmids, IL2 neurons, or both. An additional benefit is that it can also stain the excretory duct. The method allows genetic screens to be performed to identify mutants that selectively affect only one of the sensory structures or the excretory duct. Using this assay, we found that a mutation in the POU homeobox gene *unc-86* can abolish dye-filling in IL2 neurons but not amphids. Although the six IL2 (inner labial sensilla) neurons in the head also have dendrites reaching to the tip of the head with exposed endings, the IL2 neurons function is unknown. UNC-86 is not expressed in amphid sensory neurons, but *unc-86* mutants are defective in their responses to chemosensory attractants mediated by the AWC and ASE sensory neurons. Therefore the new neuron staining method provides the clue to investigate the function of IL2 neurons and the role of POU homeobox in the development of IL2 neurons.

3.3 PAPER III: THE LIM HOMEOBOX GENE *CEH-14* REGULATES NEURITE OUTGROWTH IN SELECT NEURONS

The LIM homeobox gene *ceh-14* is expressed in the AFD neurons and is required for thermotaxis behavior in *C. elegans* (Cassata et al., 2000a). In this paper, we show that *ceh-14* is also expressed in the other sensory neurons and interneurons, including the phasmid neurons (PHA, PHB, PHC, DVC, PVC, PVN, PVQ, PVT, PVW and PVR), touch neurons (ALML/R, AVM, PVM and PLML/R) and the head interneuron (ALA and BDU). Apart from the neuronal expression, *ceh-14* expression is also seen in the hypodermis and spermatheca. The expression of *ceh-14* in AFD is reduced in the *ceh-14* mutant background compared to wild type, while its expression levels in the other neurons (i.e. ALA, BDU, etc.) are not down-regulated.

In addition to severe thermotaxis defects, *ceh-14* mutant animals show no behavioral abnormalities that might be caused by defects in the sensory neurons. However, *ceh-14* mutant animals show defects in dendrite outgrowth of the phasmid neurons. Moreover, the ALA interneuron and some tail neurons show abnormal axonal outgrowth and pathfinding. *ceh-14* mutant worms have the ALA cell body at the normal position, however the axons do not run into the lateral cord, but into the ventral nerve cord (VNC) and stop just behind the posterior pharynx. Several worms have loop-shaped ALA

axons at the point where they normally turn into the lateral cord or else in the VNC. We also observed high percentages of abnormal axons of the tail neurons, aberrant projections, extra-branching and premature termination in *ceh-14* mutants. In contrast, the axonal projections from the head neurons, AFD and BDU, seemed to be unaffected. The paired-like homeobox gene *ceh-17*, which controls normal axonal outgrowth of ALA, is not regulated by *ceh-14*, nor is *ceh-14* regulated by *ceh-17*. These findings suggest that they act in separate pathways. Furthermore, *ceh-14* mutants display the body size reduction and phasmid dye-filling defects due to the abnormal development of the hypodermis. We also isolated a new allele, *ceh-14(ch2)*, in which only one LIM domain is disrupted. Analysis of this mutation shows that it behaves as a partial loss-of-function mutation.

Dissection analyses of the *ceh-14* promoter identify the regulatory regions that drive the different components of the *ceh-14* expression pattern. To beter understand its regulation, further study of the *cis*-regulatory elements of *ceh-14* is needed. To identify the downstream genes of *ceh-14* using microarrays, we did try to make the ectopic expression of the cDNA of the *ceh-14* using pan-neuronal promoter *rab-3*. However, we found that these lines were difficult to maintain. Transgenic animals arrest either in embryogenesis or L1 stage. In contrast, the LIM domain binding factor gene *ldb-1*, which is expressed in all neurons of the animal, produced a variety of defects, suggestive of important interactions in many tissues. We used an additional overexpression construct, containing only the LIM domains of CEH-14 but lacking the homeodomain. Overexpression of this construct also resulted in embryonic and L1 lethality. Taken together, these facts suggest that overexpression of CEH-14 in the nervous system titrates out interacting factors, such as LDB-1, what is the cause of the developmental defects.

3.4 PAPER IV: A REGULATORY NETWORK OF HOMEOBOX GENES IS REQUIRED FOR THE FUNCTION OF THE *CAENORHABDITIS ELEGANS* EXCRETORY CELL

We have previously described that one of the three POU homeobox genes in *C. elegans*, *ceh-6*, functions in the excretory cell which plays a key part in regulating osmotic/ionic

balance and waste elimination (Bürglin and Ruvkun, 2001). In this study, we further characterize two other homeobox genes families (*otd/Otx* and Prospero) that are required for the function of the excretory cell in *C. elegans* and the interactions among these three homeobox gene families.

The *C. elegans* genome encodes three *otd*/*Otx* orthologous. The three genes specify distinct sensory neuron identities in *C. elegans* (Lanjuin et al., 2003; Satterlee et al., 2001). Analysis of the expression pattern of *ceh-36, ceh-37* and *ttx-1* during embryogenesis using GFP reporter constructs and 2-channel 4D-microscopy indicates that the expression pattern of both *ceh-37* and *ceh-36* consist of two phases of expression, one during gastrulation, and a later one at the end of gastrulation/beginning of morphogenesis, while for the third *otd/Otx* homeobox gene, *ttx-1*, 4D analysis shows only one phase of expression after gastrulation.

The *ceh-36* deletion (*ok795*), which removes the full homeodomain and 3'UTR, results in animals with a stronger phenotype, especially in hermaphrodites. The mutant animals display reduced brood size, dead eggs, L1 arrest, and delayed development. These phenotypes contrast with the neuronal restricted phenotypes of the *ceh-36* point mutant alleles previously described (Lanjuin et al., 2003). The *ceh-36* hermaphrodites grow more slowly than males: After 72 hours, 85% of males had grown to the adult stage, while only 20% of hermaphrodites were egg-laying adults.

We also analyzed the *ceh-37* deletion mutant alleles (*tm253*, *tm254*) and *ttx-1* mutant allele (*p767*). They have no obvious defect in brood size and morphological appearance except for the neuronal defects. However, *ceh-37*(*tm253*) *ttx-1*(*p767*) double mutant animals display a severe phenotype of developmental arrest in young larval stages, with cysts in the channel of excretory cell or in the body. The necrosis of the excretory cell channels in double mutants of *otd/Otx* genes indicates an impaired development of the excretory cell. We also tested for possible redundant roles using RNAi to generate double and triple knockdowns among three *otd*/*Otx* genes. Only in the *ttx-1*/*ceh-37* double RNAi animals, we found disruptions of the excretory cell and its canals. Cysts mostly appeared in the head or the excretory cell region. Hence *ceh-37* and *ttx-1* play *a* redundant role in the excretory cell and are required for the function of the excretory cell in *C. elegans*.

Prospero gene *ceh-26* is also expressed in the excretory cell, in addition to a number of cells, primarily in the head but also in some cells in the body and the tail. 4D recording shows its wide expression in the embryo. We examined the loss of function of *ceh-26* using RNAi as well as in the recent deletion mutant allele *tm258.* With both approaches, we found a similar level of larval lethality and disruptions in the excretory cell. Therefore *ceh-26* is required for the function of the excretory cell in *C. elegans*.

We further examined the effects of knocking down *ceh-6*, *ceh-26*, and *ceh-37/ttx-1* in combination by looking at GFP reporters expressed in the excretory cell. For example, knock-down of all genes downregulates the expression of *clh-4*::GFP in the excretory cell. In contrast, expression of *sulp-4*::GFP is suppressed in *ceh-26*(*RNAi*) and *ceh-6*(*mg6*) animals, but not in *ceh-37*(*RNAi*)/*ttx-1*(*RNAi*) animals. Furthermore, only *ceh-6*(*RNAi*) completely abolishes GFP expression of *pgp-12,* but *ceh-26*(*RNAi*) and *ceh-37*(*RNAi*)/*ttx-1*(*RNAi*) only weakly reduce the level of GFP. We also investigated the regulation of the homeobox genes amongst each other using RNAi and homeobox::GFP reporter integrated lines. *ceh-6*(*RNAi*) nearly completely suppresses the GFP expression of both *ceh-37* and *ceh-26*. *ceh-26*(*RNAi*) also abolishes *ceh-37*::GFP expression. *ceh-37*/*ttx-1*(*RNAi*) of *ceh-26*::GFP showed that 42% of the arrested worms still had GFP expression, indicating a partial downregulation, possibly a regulatory feedback loop. Based on our findings, we propose a regulatory hierarchy with *ceh-6* at the top, *ceh-26* in the middle, and the *otd/Otx* genes downstream (**Figure 14**).

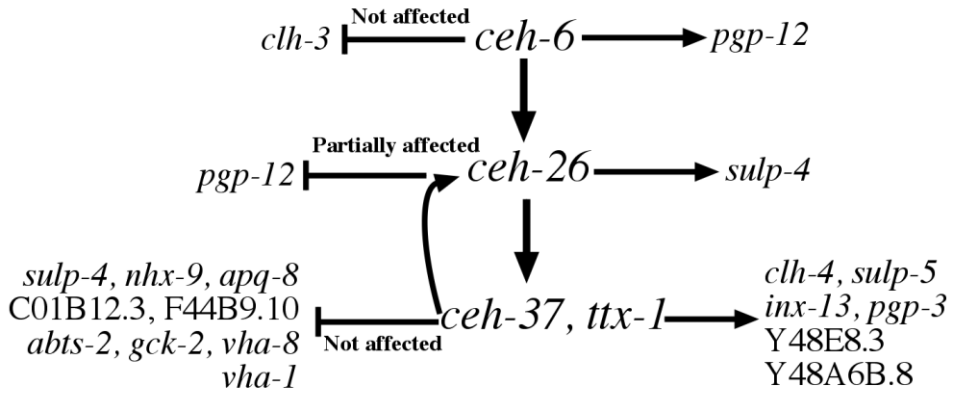

Figure 14 Model for the regulatory hierarchy in the excretory cell based on our current data in this paper.

We established that a combination of binding motifs TTTGCATAATG/CC for POU and OTD homeodomain proteins is required for *ceh-26* and *clh-4* expression in the excretory cell, using promoter deletion analyses, mutagenesis and functional studies. Our results provide further evidence for the homeobox network to exist in the excretory cell.

In conclusion, we find that two types of homeobox genes, the POU-III family and the Prospero class play a role in development of an excretory cell in a nematode *C. elegans*. Given the involvement of the homologous genes in kidney development in a wide range of species, it is intriguing to draw parallels between these processes. How well the regulatory networks and target genes are conserved in evolution is still unclear, but *C. elegans* could nevertheless provide an important model system for kidney development and function, because of its simplicity. Moreover, we have identified a cascade of homeodomain transcription factors that is essential for excretory cell function, loss of this cascade leads to lethality. Given that these homeobox genes are well conserved in evolution, we may expect that parts of this cascade are also conserved in other organisms.

4 CONCLUSIONS

1. RBD-1 plays a role in structurally coordinating pre-rRNA during ribosome biogenesis and that this function is conserved in all eukaryotes. Our data show that this gene is essential in the development of *C. elegans*.

2. IL2 dye-filling depends on salt concentration, but not osmolarity. The modified dye-filling procedure provides a reliable method to distinguish mutant alleles that stain amphids and phasmids, IL2 neurons. Using this assay, we found that a mutation in the POU homeobox gene *unc-86* abolish dye-filling in IL2 neurons, but not amphids and phasmids.

3. The LIM homeobox gene *ceh-14* and the paired-like homeobox gene *ceh-17* act in separate pathways to control normal axonal outgrowth of the ALA neuron. Overexpression of CEH-14 in the nervous system causes strong defects, most likely by titrating out interacting factors, such as LDB-1, which then causes developmental defects.

4. The *otd/Otx* homeobox genes *ceh-37* and *ttx-1* play redundant role in the excretory cell and are required for the function of the excretory cell in *C. elegans*.

5. The Prospero homeobox gene *ceh-26* is required for the function of the excretory cell in *C. elegans*.

6. A combined motif TTTGCATAATG/CC may be bound by the POU and OTD homeodomain proteins.

7. We have identified a cascade of homeodomain transcription factors comprised of *ceh-6* (POU III) at the top, *ceh-26* (Prospero) in the middle, and the *otd/Otx* genes downstream. This cascade is essential for excretory cell function, loss of it leads to lethality.

5 ACKNOWLEDGEMENTS

I would like to thank all of you who have contributed to the existence of my thesis.

Many thanks to my supervisor **Thomas Bürglin** for being truly enthusiastic about science, for introducing me to the wonderful model organism *C. elegans*, for sharing your great knowledge in bioinformatics and for the support during the past years. I am very grateful to you for the opportunity to study in the field of molecular biology. Further, I would like to thank in particular **Krai Meemon** for the fruitful collaboration, **René Prétôt** for his preliminary data on *ceh-26*, **Kiyotaka Ohkura** for help with cloning and discussions, as well as **Lois Tang** for proof reading my manuscript. The work on *ceh-14* was primarily the work of **Hiroshi Kagoshima**, and I am thankful to him for being able to contribute to this project. I am also grateful to **Lars Wieslander** and the members of his laboratory, **Petra Björk**, **Göran Baurén** and **ShaoBo Jin**, as well as **Ulf Hellman**, for the collaboration on the study of the RBD protein.

Thanks to Södertörn Högskola and Karolinska Institutet for providing financial support during my PhD study.

Thanks to Professor **Tomas Hökfelt** for making me to go to Sweden and join his excellent Lab. Thanks also to my Chinese supervisors Professor **Gong Ju** and Professor **Xu Zhang** for giving me the opportunity to study abroad.

Thanks to **Agata Smialowska** and **Lois** for using their precious time to read and correct this thesis.

Thanks to current and past members of the Bürglin lab: **Patrick Dessi**, **Lois**, **Krai**, **Kiyotaka, Jürgen Hench, Johan Henriksson**, **Johan Dethlefsen**, **Limin Hao**, **Daniela Wester**, **Krishanu Mukherjee**, **Johanna Havemann, Konstantin Cesnulevicius, LiYi Meng**, **Xiao Yu Zhao, Akram Abou-Zied**, **and Johan Koch** for scientific discussions and technical support.

Thanks to **Peter Swoboda** and members of his lab: **Agata, Jan A. Burghoorn, Ninwa Youssef, Maria Trieb, Gabi Senti, Evgeni Efimenko, Prasad Phirke, Juan Carlos**

Fierro González, Karin Furtenbach and Brian P. Piasecki for the scientific discussion, for the technical support and for sharing the material and equipment.

Thanks to colleagues in Karolinska Institutet and Södertörn Högskola: **Chiounan Shiue, Barbara Karpińska, Yongtao Xue-Franzén, Helmi Siltala, Per Kylsten, Giselbert Hauptmann, Marie Granroth, Lotta Granroth, Monica Ahlberg, Lennart Nilsson, Kristina Bergholm, Asim Syed, Ivana Bratic Hench, Claire Pujol and Qi Dai** for creating great atmospheres, help and nice times.

Thanks to my friends made in Sweden these last years especially **Mauri Kolu**, **Li Hua Guo**, **Xue Li He** and **Wei Min Zhang** for friendship, help and great fun.

Many thanks to my father **Jian Jun Tong**, mother **Su Ping Wang** and my sister **Zheng Rong Tong** for their support, patience and understanding during the best and worst of times. Thanks for everything.

Thanks to Cindy's mother, **Jun Yang Fang** for her care and support. Finally, thanks to my wonderful daughter, **Cindy Xin Yue Tong**, who was born here and has been growing up during my PhD studies. It is a great experience to be your dad, and we explored the world and made new discoveries together.

6 REFERENCES

Abate-Shen, C. (2002). Deregulated homeobox gene expression in cancer: cause or consequence? *Nat Rev Cancer* **2**, 777-85.

Acampora, D., Annino, A., Tuorto, F., Puelles, E., Lucchesi, W., Papalia, A. and Simeone, A. (2005). *Otx* genes in the evolution of the vertebrate brain. *Brain Res Bull* **66**, 410-20.

Acampora, D., Avantaggiato, V., Tuorto, F., Barone, P., Reichert, H., Finkelstein, R. and Simeone, A. (1998). Murine *Otx1* and *Drosophila otd* genes share conserved genetic functions required in invertebrate and vertebrate brain development. *Development* **125**, 1691-702.

Acampora, D., Gulisano, M., Broccoli, V. and Simeone, A. (2001). *Otx* genes in brain morphogenesis. *Prog Neurobiol* **64**, 69-95.

Acampora, D., Mazan, S., Avantaggiato, V., Barone, P., Tuorto, F., Lallemand, Y., Brulet, P. and Simeone, A. (1996). Epilepsy and brain abnormalities in mice lacking the *Otx1* gene. *Nat Genet* **14**, 218-22.

Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A. and Brulet, P. (1995). Forebrain and midbrain regions are deleted in *Otx2-/-* mutants due to a defective anterior neuroectoderm specification during gastrulation. *Development* **121**, 3279-90.

Agulnick, A. D., Taira, M., Breen, J. J., Tanaka, T., Dawid, I. B. and Westphal, H. (1996). Interactions of the LIM-domain-binding factor *Ldb1* with LIM homeodomain proteins. *Nature* **384**, 270-2.

Alberts, B. (2002). Molecular biology of the cell. New York: Garland Science.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y. and Hobert, O. (2001). A regulatory cascade of three homeobox genes, *ceh-10*, *ttx-3* and *ceh-23*, controls cell fate specification of a defined interneuron class in *C. elegans*. *Development* **128**, 1951-69.

Alvarez-Bolado, G., Rosenfeld, M. G. and Swanson, L. W. (1995). Model of forebrain regionalization based on spatiotemporal patterns of POU-III homeobox gene expression, birthdates, and morphological features. *J Comp Neurol* **355**, 237-95.

Anderson, M. G., Perkins, G. L., Chittick, P., Shrigley, R. J. and Johnson, W. A. (1995). drifter, a Drosophila POU-domain transcription factor, is required for correct differentiation and migration of tracheal cells and midline glia. *Genes Dev* **9**, 123-37.

Ang, S. L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L. and Rossant, J. (1996). A targeted mouse *Otx2* mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122**, 243-52.

Armstrong, K. R. and Chamberlin, H. M. (2010). Coordinate regulation of gene expression in the *C. elegans* excretory cell by the POU domain protein CEH-6. *Mol Genet Genomics* **283**, 73-87.

Bürglin, T. R. (2005). Homeodomain proteins. In *Encyclopedia of molecular cell biology and molecular medicine*, (ed. R. A. Meyers), pp. 179-222. Weinheim; Chichester: Wiley-VCH ; John Wiley [distributor],.

Bürglin, T. R., Finney, M., Coulson, A. and Ruvkun, G. (1989). *Caenorhabditis elegans* has scores of homoeobox-containing genes. *Nature* **341**, 239-43.
Bürglin, T. R. and Ruvkun, G. (2001). Regulation of ectodermal and excretory function by the *C. elegans* POU homeobox gene ceh-6. *Development* **128**, 779-90.
Bach, I. (2000). The LIM domain: regulation by association. *Mech Dev* **91**, 5-17.
Bach, I., Carriere, C., Ostendorff, H. P., Andersen, B. and Rosenfeld, M. G. (1997). A family of LIM domain-associated cofactors confer transcriptional synergism between LIM and *Otx* homeodomain proteins. *Genes Dev* **11**, 1370-80.
Bae, Y. K. and Barr, M. M. (2008). Sensory roles of neuronal cilia: cilia development, morphogenesis, and function in *C. elegans*. *Front Biosci* **13**, 5959-74.
Baird-Titus, J. M., Clark-Baldwin, K., Dave, V., Caperelli, C. A., Ma, J. and Rance, M. (2006). The solution structure of the native K50 Bicoid homeodomain bound to the consensus TAATCC DNA-binding site. *J Mol Biol* **356**, 1137-51.
Bargmann, C. I. (2006). Chemosensation in *C. elegans*. In *WormBook*, (ed. T. *C. e.* R. Community): WormBook.
Bargmann, C. I., Hartwieg, E. and Horvitz, H. R. (1993). Odorant-selective genes and neurons mediate olfaction in *C. elegans*. *Cell* **74**, 515-27.
Berry, K. L., Bulow, H. E., Hall, D. H. and Hobert, O. (2003). A *C. elegans* CLIC-like protein required for intracellular tube formation and maintenance. *Science* **302**, 2134-7.
Booth, H. A. and Holland, P. W. (2007). Annotation, nomenclature and evolution of four novel homeobox genes expressed in the human germ line. *Gene* **387**, 7-14.
Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71-94.
Buechner, M. (2002). Tubes and the single *C. elegans* excretory cell. *Trends Cell Biol* **12**, 479-84.
Buechner, M., Hall, D. H., Bhatt, H. and Hedgecock, E. M. (1999). Cystic canal mutants in *Caenorhabditis elegans* are defective in the apical membrane domain of the renal (excretory) cell. *Dev Biol* **214**, 227-41.
Burke, Z. and Oliver, G. (2002). Prox1 is an early specific marker for the developing liver and pancreas in the mammalian foregut endoderm. *Mech Dev* **118**, 147-55.
Burket, C. T., Higgins, C. E., Hull, L. C., Berninsone, P. M. and Ryder, E. F. (2006). The *C. elegans* gene *dig-1* encodes a giant member of the immunoglobulin superfamily that promotes fasciculation of neuronal processes. *Dev Biol* **299**, 193-205.
Casanova, J. (2007). The emergence of shape: notions from the study of the Drosophila tracheal system. *EMBO Rep* **8**, 335-9.
Cassata, G., Kagoshima, H., Andachi, Y., Kohara, Y., Durrenberger, M. B., Hall, D. H. and Bürglin, T. R. (2000a). The LIM homeobox gene *ceh-14* confers thermosensory function to the AFD neurons in *Caenorhabditis elegans*. *Neuron* **25**, 587-97.
Cassata, G., Rohrig, S., Kuhn, F., Hauri, H. P., Baumeister, R. and Bürglin, T. R. (2000b). The *Caenorhabditis elegans* Ldb/NLI/Clim orthologue *ldb-1* is required for neuronal function. *Dev Biol* **226**, 45-56.
Chang, S., Johnston, R. J., Jr. and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes Dev* **17**, 2123-37.

Charest-Marcotte, A., Dufour, C. R., Wilson, B. J., Tremblay, A. M., Eichner, L. J., Arlow, D. H., Mootha, V. K. and Giguere, V. (2010). The homeobox protein Prox1 is a negative modulator of ERR{alpha}/PGC-1{alpha} bioenergetic functions. *Genes Dev* **24**, 537-42.

Chavez, M., Landry, C., Loret, S., Muller, M., Figueroa, J., Peers, B., Rentier-Delrue, F., Rousseau, G. G., Krauskopf, M. and Martial, J. A. (1999). APH-1, a POU homeobox gene expressed in the salt gland of the crustacean Artemia franciscana. *Mech Dev* **87**, 207-12.

Chitwood, B. G. and Chitwood, M. B. (1977). Introduction to nematology. Baltimore [u.a.]: Univ. Park Pr.

Clery, A., Blatter, M. and Allain, F. H. (2008). RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290-8.

Coburn, C. M., Mori, I., Ohshima, Y. and Bargmann, C. I. (1998). A cyclic nucleotide-gated channel inhibits sensory axon outgrowth in larval and adult *Caenorhabditis elegans*: a distinct pathway for maintenance of sensory axon structure. *Development* **125**, 249-58.

Collet, J., Spike, C. A., Lundquist, E. A., Shaw, J. E. and Herman, R. K. (1998). Analysis of *osm-6*, a gene that affects sensory cilium structure and sensory neuron function in *Caenorhabditis elegans*. *Genetics* **148**, 187-200.

Consortium, T. *C. e.* S. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012-8.

Cui, W., Tomarev, S. I., Piatigorsky, J., Chepelinsky, A. B. and Duncan, M. K. (2004). Mafs, Prox1, and Pax6 can regulate chicken betaB1-crystallin gene expression. *J Biol Chem* **279**, 11088-95.

de Bono, M. (2003). Molecular approaches to aggregation behavior and social attachment. *J Neurobiol* **54**, 78-92.

de Celis, J. F., Llimargas, M. and Casanova, J. (1995). Ventral veinless, the gene encoding the Cf1a transcription factor, links positional information and cell differentiation during embryonic and imaginal development in Drosophila melanogaster. *Development* **121**, 3405-16.

Derelle, R., Lopez, P., Le Guyader, H. and Manuel, M. (2007). Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. *Evol Dev* **9**, 212-9.

Ding, J., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R. and Xu, R. M. (1999). Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* **13**, 1102-15.

Dreyfuss, G., Kim, V. N. and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* **3**, 195-205.

Duboule, D. (1994). Guidebook to the homeobox genes. Oxford; New York: Oxford University Press.

Dudas, J., Mansuroglu, T., Moriconi, F., Haller, F., Wilting, J., Lorf, T., Fuzesi, L. and Ramadori, G. (2008). Altered regulation of Prox1-gene-expression in liver tumors. *BMC Cancer* **8**, 92.

Durbin, R. M. (1987). Studies on the development and organisation of the nervous system of *Caenorhabditis elegans*: University of Cambridge.

Epstein, H. F., Shakes, D. C. and American Society for Cell, B. (1995). *Caenorhabditis elegans* : modern biological analysis of an organism. San Diego: Academic Press.
Etchberger, J. F., Flowers, E. B., Poole, R. J., Bashllari, E. and Hobert, O. (2009). Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in *C. elegans*. *Development* **136**, 147-60.
Evans (ed.), T. C. (2006). Transformation and microinjection. In *WormBook*, (ed. T. *C. e.* R. Community): WormBook.
Feng, Q., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P. and Zhang, Y. (2002). Identification and functional characterization of the p66/p68 components of the MeCP1 complex. *Mol Cell Biol* **22**, 536-46.
Fields, S. and Johnston, M. (2005). Cell biology. Whither model organism research? *Science* **307**, 1885-6.
Finkelstein, R. and Perrimon, N. (1990). The *orthodenticle* gene is regulated by bicoid and torso and specifies *Drosophila* head development. *Nature* **346**, 485-8.
Finkelstein, R., Smouse, D., Capaci, T. M., Spradling, A. C. and Perrimon, N. (1990). The orthodenticle gene encodes a novel homeo domain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. *Genes Dev* **4**, 1516-27.
Finney, M. and Ruvkun, G. (1990). The unc-86 gene product couples cell lineage and cell identity in C. elegans. *Cell* **63**, 895-905.
Finney, M., Ruvkun, G. and Horvitz, H. R. (1988). The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors. *Cell* **55**, 757-69.
Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-11.
Fox, M. A. (1985). The case for animal experimentation. Berkeley: University of California Press.
Fujita, M., Hawkinson, D., King, K. V., Hall, D. H., Sakamoto, H. and Buechner, M. (2003). The role of the ELAV homologue EXC-7 in the development of the *Caenorhabditis elegans* excretory canals. *Dev Biol* **256**, 290-301.
Galliot, B., de Vargas, C. and Miller, D. (1999). Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class. *Dev Genes Evol* **209**, 186-97.
Gehring, W. J. (1994). Guidebook to the Homeobox Genes.
Gehring, W. J. (1998). Master Control Genes in Development and Evolution: The Homeobox Story. New Haven,USA: Yale University Press.
Greenstein, D., Hird, S., Plasterk, R. H., Andachi, Y., Kohara, Y., Wang, B., Finney, M. and Ruvkun, G. (1994). Targeted mutations in the Caenorhabditis elegans POU homeo box gene ceh-18 cause defects in oocyte cell cycle arrest, gonad migration, and epidermal differentiation. *Genes Dev* **8**, 1935-48.
Hahn-Windgassen, A. and Van Gilst, M. R. (2009). The *Caenorhabditis elegans* HNF4alpha Homolog, NHR-31, mediates excretory tube growth and function through coordinate regulation of the vacuolar ATPase. *PLoS Genet* **5**, e1000553.

Hall, D. H., Lints, R. and Altun, Z. (2006). Nematode neurons: anatomy and anatomical methods in *Caenorhabditis elegans*. *Int Rev Neurobiol* **69**, 1-35.

Hassan, B., Li, L., Bremer, K. A., Chang, W., Pinsonneault, J. and Vaessin, H. (1997). Prospero is a panneural transcription factor that modulates homeodomain protein activity. *Proc Natl Acad Sci U S A* **94**, 10991-6.

He, X., Treacy, M. N., Simmons, D. M., Ingraham, H. A., Swanson, L. W. and Rosenfeld, M. G. (1989). Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* **340**, 35-41.

Hedgecock, E. M., Culotti, J. G., Thomson, J. N. and Perkins, L. A. (1985). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. *Dev Biol* **111**, 158-70.

Hedgecock, E. M. and Russell, R. L. (1975). Normal and mutant thermotaxis in the nematode Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **72**, 4061-5.

Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D., Sharp, P. A., Ingraham, H. A., Rosenfeld, M. G., Finney, M., Ruvkun, G. et al. (1988). The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. *Genes Dev* **2**, 1513-6.

Hirata, J., Nakagoshi, H., Nabeshima, Y. and Matsuzaki, F. (1995). Asymmetric segregation of the homeodomain protein Prospero during *Drosophila* development. *Nature* **377**, 627-30.

Hirth, F., Therianos, S., Loop, T., Gehring, W. J., Reichert, H. and Furukubo-Tokunaga, K. (1995). Developmental defects in brain segmentation caused by mutations of the homeobox genes orthodenticle and empty spiracles in *Drosophila*. *Neuron* **15**, 769-78.

Hobert, O., D'Alberti, T., Liu, Y. and Ruvkun, G. (1998). Control of neural development and function in a thermoregulatory network by the LIM homeobox gene *lin-11*. *J Neurosci* **18**, 2084-96.

Hobert, O. and Westphal, H. (2000). Functions of LIM-homeobox genes. *Trends Genet* **16**, 75-83.

Holland, P. W., Booth, H. A. and Bruford, E. A. (2007). Classification and nomenclature of all human homeobox genes. *BMC Biol* **5**, 47.

Hutter, H. and Schnabel, R. (1994). *glp-1* and inductions establishing embryonic axes in *C. elegans*. *Development* **120**, 2051-64.

Inglis, P. N., Ou, G., Leroux, M. R. and Scholey, J. M. (2007). The sensory cilia of *Caenorhabditis elegans*. In *WormBook*, (ed. T. C. e. R. Community): WormBook.

Ingraham, H. A., Flynn, S. E., Voss, J. W., Albert, V. R., Kapiloff, M. S., Wilson, L. and Rosenfeld, M. G. (1990). The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. *Cell* **61**, 1021-33.

Johnson, N. C., Dillard, M. E., Baluk, P., McDonald, D. M., Harvey, N. L., Frase, S. L. and Oliver, G. (2008). Lymphatic endothelial cell identity is reversible and its maintenance requires Prox1 activity. *Genes Dev* **22**, 3282-91.

Jones, S. J. and Baillie, D. L. (1995). Characterization of the *let-653* gene in Caenorhabditis elegans. *Mol Gen Genet* **248**, 719-26.

Jung, A. C., Denholm, B., Skaer, H. and Affolter, M. (2005). Renal tubule development in *Drosophila*: a closer look at the cellular level. *J Am Soc Nephrol* **16**, 322-8.

Jurata, L. W., Kenny, D. A. and Gill, G. N. (1996). Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. *Proc Natl Acad Sci U S A* **93**, 11693-8.

Kadrmas, J. L. and Beckerle, M. C. (2004). The LIM domain: from the cytoskeleton to the nucleus. *Nat Rev Mol Cell Biol* **5**, 920-31.

Kim, J. B., Sebastiano, V., Wu, G., Arauzo-Bravo, M. J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrich, M., van den Boom, D. et al. (2009). Oct4-induced pluripotency in adult neural stem cells. *Cell* **136**, 411-9.

Kim, K., Kim, R. and Sengupta, P. (2010). The HMX/NKX homeodomain protein MLS-2 specifies the identity of the AWC sensory neuron type via regulation of the *ceh-36* Otx gene in *C. elegans*. *Development* **137**, 963-74.

Klambt, C., Jacobs, J. R. and Goodman, C. S. (1991). The midline of the *Drosophila* central nervous system: a model for the genetic analysis of cell fate, cell migration, and growth cone guidance. *Cell* **64**, 801-15.

Koga, M. and Ohshima, Y. (2004). The *C. elegans ceh-36* gene encodes a putative homemodomain transcription factor involved in chemosensory functions of ASE and AWC neurons. *J Mol Biol* **336**, 579-87.

Kohler, R. E. (1994). Lords of the fly : *Drosophila* genetics and the experimental life. Chicago: University of Chicago Press.

Lan, L., Liu, M., Liu, Y., Zhang, W., Xue, J., Xue, Z. and He, R. (2006). Expression of qBrn-1, a new member of the POU gene family, in the early developing nervous system and embryonic kidney. *Dev Dyn* **235**, 1107-14.

Lanjuin, A., VanHoven, M. K., Bargmann, C. I., Thompson, J. K. and Sengupta, P. (2003). Otx/otd homeobox genes specify distinct sensory neuron identities in *C. elegans*. *Dev Cell* **5**, 621-33.

Lee, M.-H. and Schedl, T. (2006). RNA-binding proteins. In *WormBook*, (ed. T. *C. e.* R. Community): WormBook.

Lee, S., Kang, J., Yoo, J., Ganesan, S. K., Cook, S. C., Aguilar, B., Ramu, S., Lee, J. and Hong, Y. K. (2009). Prox1 physically and functionally interacts with COUP-TFII to specify lymphatic endothelial cell fate. *Blood* **113**, 1856-9.

Lesch, B. J., Gehrke, A. R., Bulyk, M. L. and Bargmann, C. I. (2009). Transcriptional regulation and stabilization of left-right neuronal identity in *C. elegans*. *Genes Dev* **23**, 345-58.

Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-70.

Lichtsteiner, S. and Tjian, R. (1995). Synergistic activation of transcription by UNC-86 and MEC-3 in Caenorhabditis elegans embryo extracts. *EMBO J* **14**, 3937-45.

Llimargas, M. and Casanova, J. (1997). ventral veinless, a POU domain transcription factor, regulates different transduction pathways required for tracheal branching in Drosophila. *Development* **124**, 3273-81.

Loria, P. M., Duke, A., Rand, J. B. and Hobert, O. (2003). Two neuronal, nuclear-localized RNA binding proteins involved in synaptic transmission. *Curr Biol* **13**, 1317-23.

Mah, A. K., Armstrong, K. R., Chew, D. S., Chu, J. S., Tu, D. K., Johnsen, R. C., Chen, N., Chamberlin, H. M. and Baillie, D. L. (2007). Transcriptional regulation of AQP-8, a *Caenorhabditis elegans* aquaporin exclusively expressed in the excretory system, by the POU homeobox transcription factor CEH-6. *J Biol Chem* **282**, 28074-86.

Mah, A. K., Tu, D. K., Johnsen, R. C., Chu, J. S., Chen, N. and Baillie, D. L. (2010). Characterization of the octamer, a cis-regulatory element that modulates excretory cell gene-expression in *Caenorhabditis elegans*. *BMC Mol Biol* **11**, 19.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. et al. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**, 835-9.

Mao, C. A., Gan, L. and Klein, W. H. (1994). Multiple *Otx* binding sites required for expression of the *Strongylocentrotus purpuratus Spec2a* gene. *Dev Biol* **165**, 229-42.

Maris, C., Dominguez, C. and Allain, F. H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**, 2118-31.

Matsuo, I., Kuratani, S., Kimura, C., Takeda, N. and Aizawa, S. (1995). Mouse *Otx2* functions in the formation and patterning of rostral head. *Genes Dev* **9**, 2646-58.

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. and Gehring, W. J. (1984). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell* **37**, 403-8.

Milan, M. and Cohen, S. M. (1999). Regulation of LIM homeodomain activity in vivo: a tetramer of dLDB and apterous confers activity and capacity for regulation by dLMO. *Mol Cell* **4**, 267-73.

Mori, I. and Ohshima, Y. (1995). Neural regulation of thermotaxis in *Caenorhabditis elegans*. *Nature* **376**, 344-8.

Mukhopadhyay, M., Teufel, A., Yamashita, T., Agulnick, A. D., Chen, L., Downs, K. M., Schindler, A., Grinberg, A., Huang, S. P., Dorward, D. et al. (2003). Functional ablation of the mouse Ldb1 gene results in severe patterning defects during gastrulation. *Development* **130**, 495-505.

Nagao, T., Leuzinger, S., Acampora, D., Simeone, A., Finkelstein, R., Reichert, H. and Furukubo-Tokunaga, K. (1998). Developmental rescue of *Drosophila* cephalic defects by the human *Otx* genes. *Proc Natl Acad Sci U S A* **95**, 3737-42.

Nelson, F. K., Albert, P. S. and Riddle, D. L. (1983). Fine structure of the *Caenorhabditis elegans* secretory-excretory system. *J Ultrastruct Res* **82**, 156-71.

Nelson, F. K. and Riddle, D. L. (1984). Functional study of the *Caenorhabditis elegans* secretory-excretory system using laser microsurgery. *J Exp Zool* **231**, 45-56.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H. and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-91.

Nishijima, I. and Ohtoshi, A. (2006). Characterization of a novel prospero-related homeobox gene, Prox2. *Mol Genet Genomics* **275**, 471-8.

Perkins, L. A., Hedgecock, E. M., Thomson, J. N. and Culotti, J. G. (1986). Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Dev Biol* **117**, 456-87.

Pierani, A., Heguy, A., Fujii, H. and Roeder, R. G. (1990). Activation of octamer-containing promoters by either octamer-binding transcription factor 1 (OTF-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. *Mol Cell Biol* **10**, 6204-15.

Poulin, G., Dong, Y., Fraser, A. G., Hopper, N. A. and Ahringer, J. (2005). Chromatin regulation and sumoylation in the inhibition of Ras-induced vulval development in *Caenorhabditis elegans*. *EMBO J* **24**, 2613-23.

Qian, Y. Q., Billeter, M., Otting, G., Muller, M., Gehring, W. J. and Wuthrich, K. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **59**, 573-80.

Qin, J., Gao, D. M., Jiang, Q. F., Zhou, Q., Kong, Y. Y., Wang, Y. and Xie, Y. H. (2004). Prospero-related homeobox (Prox1) is a corepressor of human liver receptor homolog-1 and suppresses the transcription of the cholesterol 7-alpha-hydroxylase gene. *Mol Endocrinol* **18**, 2424-39.

Ranade, S. S., Yang-Zhou, D., Kong, S. W., McDonald, E. C., Cook, T. A. and Pignoni, F. (2008). Analysis of the *Otd*-dependent transcriptome supports the evolutionary conservation of CRX/OTX/OTD functions in flies and vertebrates. *Dev Biol* **315**, 521-34.

Rankin, C. H. (2002). From gene to identified neuron to behaviour in *Caenorhabditis elegans*. *Nat Rev Genet* **3**, 622-30.

Remy, J. J. and Hobert, O. (2005). An interneuronal chemoreceptor required for olfactory imprinting in *C. elegans*. *Science* **309**, 787-90.

Robinson-Rechavi, M., Maina, C. V., Gissendanner, C. R., Laudet, V. and Sluder, A. (2005). Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. *J Mol Evol* **60**, 577-86.

Rodriguez-Niedenfuhr, M., Papoutsi, M., Christ, B., Nicolaides, K. H., von Kaisenberg, C. S., Tomarev, S. I. and Wilting, J. (2001). Prox1 is a marker of ectodermal placodes, endodermal compartments, lymphatic endothelium and lymphangioblasts. *Anat Embryol (Berl)* **204**, 399-406.

Rose, K. L., Winfrey, V. P., Hoffman, L. H., Hall, D. H., Furuta, T. and Greenstein, D. (1997). The POU gene ceh-18 promotes gonadal sheath cell differentiation and function required for meiotic maturation and ovulation in Caenorhabditis elegans. *Dev Biol* **192**, 59-77.

Ruvkun, G. and Hobert, O. (1998). The taxonomy of developmental control in *Caenorhabditis elegans*. *Science* **282**, 2033-41.

Ryan, A. K. and Rosenfeld, M. G. (1997). POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* **11**, 1207-25.

Satterlee, J. S., Sasakura, H., Kuhara, A., Berkeley, M., Mori, I. and Sengupta, P. (2001). Specification of thermosensory neuron fate in *C. elegans* requires *ttx-1*, a homolog of *otd/Otx*. *Neuron* **31**, 943-56.

Scott, M. P. and Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of *Drosophila*. *Proc Natl Acad Sci U S A* **81**, 4115-9.

Shaham (ed.), S. (2006). Methods in cell biology. In *WormBook*, (ed. T. *C. e.* R. Community): WormBook.

Shamoo, Y., Abdul-Manan, N. and Williams, K. R. (1995). Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res* **23**, 725-8.

Sharman, A. C. and Brand, M. (1998). Evolution and homology of the nervous system: cross-phylum rescues of *otd/Otx* genes. *Trends Genet* **14**, 211-4.

Simeone, A., Acampora, D., Gulisano, M., Stornaiuolo, A. and Boncinelli, E. (1992). Nested expression domains of four homeobox genes in developing rostral brain. *Nature* **358**, 687-90.

Song, K. H., Li, T. and Chiang, J. Y. (2006). A Prospero-related homeodomain protein is a novel co-regulator of hepatocyte nuclear factor 4alpha that regulates the cholesterol 7alpha-hydroxylase gene. *J Biol Chem* **281**, 10081-8.

Sosa-Pineda, B., Wigle, J. T. and Oliver, G. (2000). Hepatocyte migration during liver development requires Prox1. *Nat Genet* **25**, 254-5.

Spana, E. P. and Doe, C. Q. (1995). The prospero transcription factor is asymmetrically localized to the cell cortex during neuroblast mitosis in *Drosophila*. *Development* **121**, 3187-95.

Spaniol, P., Bornmann, C., Hauptmann, G. and Gerster, T. (1996). Class III POU genes of zebrafish are predominantly expressed in the central nervous system. *Nucleic Acids Res* **24**, 4874-81.

Starich, T. A., Herman, R. K., Kari, C. K., Yeh, W. H., Schackwitz, W. S., Schuyler, M. W., Collet, J., Thomas, J. H. and Riddle, D. L. (1995). Mutations affecting the chemosensory neurons of *Caenorhabditis elegans*. *Genetics* **139**, 171-88.

Staudt, L. M., Singh, H., Sen, R., Wirth, T., Sharp, P. A. and Baltimore, D. (1986). A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature* **323**, 640-3.

Stiernagle, T. (2006). Maintenance of *C. elegans*. In *WormBook*, (ed. T. C. e. R. Community): WormBook.

Sulston, J. E. (1983). Neuronal cell lineages in the nematode *Caenorhabditis elegans*. *Cold Spring Harb Symp Quant Biol* **48 Pt 2**, 443-52.

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* **56**, 110-56.

Sulston, J. E., Schierenberg, E., White, J. G. and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* **100**, 64-119.

Tahayato, A., Sonneville, R., Pichaud, F., Wernet, M. F., Papatsenko, D., Beaufils, P., Cook, T. and Desplan, C. (2003). *Otd/Crx*, a dual regulator for the specification of ommatidia subtypes in the Drosophila retina. *Dev Cell* **5**, 391-402.

Takatori, N., Butts, T., Candiani, S., Pestarino, M., Ferrier, D. E., Saiga, H. and Holland, P. W. (2008). Comprehensive survey and classification of homeobox genes in the genome of amphioxus, Branchiostoma floridae. *Dev Genes Evol* **218**, 579-90.

Tong, X. and Buechner, M. (2008). CRIP homologues maintain apical cytoskeleton to regulate tubule size in *C. elegans*. *Dev Biol* **317**, 225-33.

Turner, F. R. and Mahowald, A. P. (1979). Scanning electron microscopy of *Drosophila* melanogaster embryogenesis. III. Formation of the head and caudal segments. *Dev Biol* **68**, 96-109.

Verrijzer, C. P., Kal, A. J. and van der Vliet, P. C. (1990). The oct-1 homeo domain contacts only part of the octamer sequence and full oct-1 DNA-binding activity requires the POU-specific domain. *Genes Dev* **4**, 1964-74.

Wade, P. A., Gegonne, A., Jones, P. L., Ballestar, E., Aubry, F. and Wolffe, A. P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat Genet* **23**, 62-6.

Wang, X. and Chamberlin, H. M. (2002). Multiple regulatory changes contribute to the evolution of the *Caenorhabditis lin-48* ovo gene. *Genes Dev* **16**, 2345-9.

Wang, X. and Chamberlin, H. M. (2004). Evolutionary innovation of the excretory system in *Caenorhabditis elegans*. *Nat Genet* **36**, 231-2.

Ward, S., Thomson, N., White, J. G. and Brenner, S. (1975). Electron microscopical reconstruction of the anterior sensory anatomy of the nematode *Caenorhabditis elegans*. *J Comp Neurol* **160**, 313-37.

Ware, R. W. (1975). Three-dimensional reconstruction from serial sections. *Int Rev Cytol* **40**, 325-440.

Wegner, M., Drolet, D. W. and Rosenfeld, M. G. (1993). Regulation of JC virus by the POU-domain transcription factor Tst-1: implications for progressive multifocal leukoencephalopathy. *Proc Natl Acad Sci U S A* **90**, 4743-7.

White, J. G. S., E.; Thomson, J. N.; Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences,* **314**, 340.

Wieschaus, E., Perrimon, N. and Finkelstein, R. (1992). orthodenticle activity is required for the development of medial structures in the larval and adult epidermis of *Drosophila*. *Development* **115**, 801-11.

Wigle, J. T. and Oliver, G. (1999). Prox1 function is required for the development of the murine lymphatic system. *Cell* **98**, 769-78.

Witta, S. E., Agarwal, V. R. and Sato, S. M. (1995). XIPOU 2, a noggin-inducible gene, has direct neuralizing activity. *Development* **121**, 721-30.

Wolberger, C. (1996). Homeodomain interactions. *Curr Opin Struct Biol* **6**, 62-8.

Wood, W. B. (1988). Introduction to *C. elegans* biology. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Yamazaki, T., Yoshimatsu, Y., Morishita, Y., Miyazono, K. and Watabe, T. (2009). COUP-TFII regulates the functions of Prox1 in lymphatic endothelial cells through direct interaction. *Genes Cells* **14**, 425-34.

Younossi-Hartenstein, A., Green, P., Liaw, G. J., Rudolph, K., Lengyel, J. and Hartenstein, V. (1997). Control of early neurogenesis of the *Drosophila* brain by the head gap genes tll, otd, ems, and btd. *Dev Biol* **182**, 270-83.

Zhao, Z., Fang, L., Chen, N., Johnsen, R. C., Stein, L. and Baillie, D. L. (2005). Distinct regulatory elements mediate similar expression patterns in the excretory cell of *Caenorhabditis elegans*. *J Biol Chem* **280**, 38787-94.

Zinovieva, R. D., Duncan, M. K., Johnson, T. R., Torres, R., Polymeropoulos, M. H. and Tomarev, S. I. (1996). Structure and chromosomal localization of the human homeobox gene Prox 1. *Genomics* **35**, 517-22.

Zipperlen, P., Fraser, A. G., Kamath, R. S., Martinez-Campos, M. and Ahringer, J. (2001). Roles for 147 embryonic lethal genes on *C.elegans* chromosome I identified by RNA interference and video microscopy. *EMBO J* **20**, 3984-92.